Structural studies on Galectin-7 and Sirtuins

Ramaswamy, Sneha

Award date:
2015

Awarding institution:
University of Bath

Structural Studies on Galectin-7 and Sirtuins

A thesis submitted for the degree of

Doctor of Philosophy

By

Sneha Ramaswamy

Department of Biology and Biochemistry

University of Bath

October 2014

Dedicated to,
my lovely parents, Mrs. Radha Ramaswamy and Dr. L. Ramaswamy, and brother, Sharath Ramaswamy
For your endless support and affection

ACKNOWLEDGEMENTS

My most sincere thanks to my supervisor, Prof. K. Ravi Acharya, for giving me the opportunity to pursue this PhD under his guidance. I would like to thank him for introducing me to structural biology, especially macromolecular crystallography and for sharing his enthusiasm for this subject. I will remain eternally grateful for his patience and support all through this journey.

I would like to thank our collaborators, Dr. Barry Steele and his group who provided us with the dendrimers used in this thesis. I would also like to acknowledge the beam-line scientists at the Diamond Light Source, U.K. where the X-ray diffraction data for hGal-7 in complex with the multivalent dendrons were collected. I thank Dr. Jean van de Elsen for his help in the interpretation of dynamic light scattering experiments. I would like to extend my appreciation to Dr. Matthew Lloyd and Dr. Paul De Bank, for their guidance in performing sirtuin assays. I am indebted to the University of Bath for the financial support through their graduate school scholarship and the Department of Biology and Biochemistry for their facilities.

I would like to thank Dr. Ami Miller, Dr. Shalini Iyer, Dr. Geoffrey Masuyer and Mr. William Bradshaw, for their useful discussion on structural biology, crystallisation and processing of diffraction data. A special thanks to Ami and Will who helped proof read this thesis. My sincere gratitude to past lab members Dr. Nethaji Thiyagarajan and Dr. Elizabeth Clark for their help and guidance. I am very grateful to all members of Lab 0.34, especially Abigail, Charlotte, Stefanie, Saima and Tram, for their support all through my PhD.

My heartfelt gratitude to my parents and my brother, Sharath, for always believing in me, standing by me and helping me achieve all I wished for. I wish to specially thank my father, Dr. L. Ramaswamy, who has been my biggest source of inspiration and motivation to pursue a career in research. Without my mother's wisdom and constant support at all times, I would not be here today. Living away from my family has been the toughest part of studying abroad.

Finally, I would like to thank my soon to be husband, Abishek, without whom I would not have survived through this journey. I truly appreciate you for just 'being' there.

ABSTRACT

Over 100 years ago, X-ray crystallography gave scientists a window to the atomic world with varied applications in biology, chemistry and physics among other subjects. Macromolecular crystallography is now considered an essential tool for solving the three-dimensional structure of proteins and understanding their physiological role at the atomic level. As crystal growth remains a bottleneck in crystallography, various other techniques are often employed to help understand the protein structure and function. These methods range from simple analysis of the protein sequence to experiments such as dynamic light scattering, isothermal titration calorimetry, activity assays, and analytical ultracentrifugation. The additional knowledge gained about proteins from these methods can then assist in the modification of the protein to facilitate its crystallisation. The structural biology of proteins belonging to two diverse families; Galectins and Sirtuins, both involved in the regulation of cancer, was studied in this thesis.

Galectins are evolutionarily conserved and ubiquitously present animal lectins with a high affinity for β-galactose containing oligosaccharides. To date, 15 mammalian galectins have been identified. Their involvement in cell–cell and cell–matrix interactions has highlighted their importance in signal transduction and other intracellular processes. Human Galectin-7 (hGal-7) is a 16 kDa prototype galectin which is involved in the stimulation and development of cancer. The crystal structure of native hGal-7 and its complex with galactose and lactose have been reported. In this study, cross-linking of hGal-7 by glycodendrons and the resulting clustering and lattice formation have been studied. For this purpose, the high resolution X-ray structures of hGal-7 in complex with carbohydrate-based multivalent dendrons have been elucidated and analysed. Also discussed in this thesis are preliminary binding affinity results obtained using isothermal calorimetry. Supramolecular assembly formation was also assessed using dynamic light scattering. These experiments reveal how multivalent glycodendendrons interact with and form cross-links with hGal-7 molecules. Understanding how these dendrimeric compounds

interact with hGal-7 would help in the design of new tools to investigate the recognition of multivalent carbohydrates by lectins and their resulting role in aggregation processes in tumour embolisation and survival.

Sirtuins are NAD^+-dependent deacylases that are involved in the regulation of diverse biological functions such as ageing, metabolism and stress resistance, in normal cellular physiology. Their role in ageing and ageing-related diseases, including cancer and neurodegenerative diseases among others, has received much attention and sirtuins have been extensively studied to help in extension of human lifespan. Seven members of the sirtuin family (SIRT1-7) are known, which are diversely sub-localised in the cell. A myriad of questions regarding their deacylation activity, their interplay and their role in various diseases still remain unanswered. In this study, structural biology techniques have been used to understand the role of SIRT1, SIRT2 and SIRT7. The cloning, expression, purification and crystallisation of these sirtuins are presented in this thesis. Various supporting techniques used to confirm the identity and activity of the proteins are also discussed. A brief discussion of the methods that can be employed to overcome various barriers in structural biology is also presented in this thesis. Elucidating the structure of full length sirtuins would help in the development of highly selective modulators of sirtuins to aid in the understanding of their role in ageing and ageing-related diseases.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction to Galectins

Chapter 2. Introduction to Sirtuins

Chapter 3: Materials and Methods

Chapter 4: Multivalent galactose based dendrimer recognition by human galectin-7

Chapter 5. Structural studies of Sirtuins

SIRT2

SIRT7

SIRT1

Discussion

LIST OF TABLES

Chapter 1: Introduction to Galectins.

Chapter 2: Introduction to Sirtuins.

Chapter 3: Materials and Methods

Chapter 4: Multivalent galactose based dendrimers by human galectin -7.

SIRT1

LIST OF ABBREVIATIONS

Å	angstrom (10^{-10} m)
Ab	antibody
ADP	adenosine diphosphate
ADPr	adenosine diphosphate ribose
AI	autoinduction
AIEX	anion exchange chromatography
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
CIEX	cation exchange chromatography
CRD	carbohydrate recognition domain
DNA	deoxyribonucleic acid
dNTP	deoxynucleotides triphosphate
ECM	extra-cellular matrix
EDTA	Ethylenediaminetetraacetatic acid
FPLC	fast performance liquid chromatography
hGal	Human galectin
GST	glutathione-S-transferase
h	hour
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
His tag	$(polyhistidine)_6$-tag
HRP	horse radish peroxidase
IL	Interleukin
IMAC	immobilised metal chelating chromatography
IPTG	isopropyl-D-β-galactopyranoside
kb	kilobase
kDa	kilodalton
K_d	dissociation constant
L	Litre
LB	Luria-Bertani
μ	micro (10^{-6})
m	milli (10^{-3})
M	molar
MALDI	matrix assisted laser desorption ionisation
min	minute
MPD	2-methyl-pentane-2,4-diol
MS	mass spectrometry
NAD^+	Nicotinamide adenine dinucleotide (oxidised form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (reduced form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NMR	nuclear magnetic resonance spectrometry
n	nano (10^{-9})
NCS	non-crystallographic symmetry

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethyleneglycol
pI	isoelectric point
rpm	revolutions per minute
rmsd	root mean square distance
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
RFU	Relative fluorescence units
s	seconds
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SIRT	sirtuin
Tris	Tris(hydroxymethyl)aminomethane
TEMED	N,N,N,N-tetramethylendiamine
TCEP	Tri(carboxyethyl) phosphine
TB	terrific broth
TBS	Tris buffer system
TNF-α	tumour necrosis factor alpha
UV	ultraviolet
vol	volumes
(v/v)	volume per volume
(w/v)	weight per volume

Amino acids, one and three letter codes

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Aspartic acid	Asp	D
Asparagine	Asn	N
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1. INTRODUCTION TO GALECTINS

1.1. Overview

Galectins are a family of proteins belonging to the lectin superfamily. They behave as signalling molecules by binding to β-galactose containing glycoconjugates at their carbohydrate recognition domains. This chapter describes the diversity of lectins and more specifically the galectin family. It also presents an overview of the biological activities and structural similarities of galectins and the functionality of human galectin–7 (hGal-7). This is followed by a brief description of carbohydrates and dendrimers, and finally the aims and objectives of this project.

1.1.1. The history of lectins

The 20th century witnessed many important scientific breakthroughs, one of which was understanding the occurrence of erythrocyte agglutinating proteins in nature (Franz, 1988). The ability of these proteins to agglutinate other proteins and their high level of sugar specificity came to light in the 1960s. The potentiality of plant agglutinins to distinguish between erythrocytes of different blood types led Boyd and Shapleigh to name them lectins (from the Latin *legere*, "to select") (Boyd and Shapleigh, 1954). This term was generalised by Sharon and Lis to include all sugar-specific agglutinins of non-immune origin, irrespective of source and blood type specificity (Sharon and Lis, 1972).

Boyd and Shapleigh's work was then followed by the two major discoveries which led to understanding the role of lectins in mitogenic stimulation of lymphocytes and agglutination of cancer cells. Peter C. Nowell found that phytohaemagglutinin (PHA), the lectin of red kidney beans (*Phaseolus vulgaris*), possessed the ability to stimulate lymphocytes to undergo mitosis *i.e.* PHA is mitogenic (Nowell, 1960). The second discovery, made by Joseph C. Aub demonstrated that wheat germ agglutinin (WGA) has the ability to preferentially agglutinate malignant cells (Aub *et al.*, 1965, Aub *et al.*, 1963).

By the 1970s the presence of haemagglutinins had mainly been reported in plants (Sharon and Lis, 2004). Among animal lectins, only haemagglutinins in eels, snails and the horseshoe crabs had been isolated and characterised by this

time. In 1952 Watkins and Morgan showed the specificity of the eel haemagglutinin for L-fucose (a sugar) (Watkins and Morgan, 1952). The first mammalian lectin, the galactose-specific hepatic asialo-glycoprotein receptor, was isolated in 1974 (Hudgin *et al.*, 1974, Stockert *et al.*, 1974) and resulted in our understanding of the mechanisms that control the lifetime of glycoproteins in blood circulation.

Following the discovery of the hepatic asialo-glycoprotein receptor, a C-type lectin, the first member of β-galactose specific lectins from the electric eel termed, electrolectin was identified in 1975 (Teichberg *et al.*, 1975). This protein required the inclusion of β-mercaptoethanol in isolation buffers to maintain its activity, which suggested the presence of one or more reduced sulfhydryl cysteine residues. These lectins were formally termed S-type lectins due to their sulphydryl dependency. This family of proteins were later termed galectins (Barondes *et al.*, 1994a) of which 15 members have now been characterised. The first mammalian galectin, now known as galectin-1, was isolated in 1976 from chick muscle and extracts of calf heart and lung (de Waard *et al.*, 1976) .

1.1.2. Animal lectins

Lectins are non-enzymatic, sugar binding proteins. They are ubiquitously present in animals, plants and microorganisms. Their specificity and ability to discriminate among the myriad of complex carbohydrates found on cell surfaces, extracellular matrices and attached to soluble glycoproteins, makes them an important tool in the structural and functional analysis of animal cell glycoconjugates (Drickamer and Taylor, 1993). These complex saccharides have both structural and non-structural roles. The non-structural role of sugars requires the participation of sugar-binding proteins, which is ascribed to a single protein module within the lectin polypeptide. This module is called the carbohydrate recognition domain (CRD) (Drickamer and Taylor, 1998). Lectins serve as a bridge between cells and their environment through their recognition of these diverse complex oligosaccharides. Their main function is in the mediation of cell-cell adhesion, intracellular routing of glycoproteins and promotion of cell-cell interactions contributing to signal transductions (Liu and

Rabinovich, 2005, Vasta, 2009). In addition to recognising carbohydrates endogenous to the animal cell, they also recognise those presented by microbial invaders and contribute to the initiation of the host defence. It is believed that lectins mediate the interactions by decoding the information present in saccharides (Drickamer and Taylor, 1993). In 1991 Gabius demonstrated the abundance and variety of animal lectins through the large number of polypeptides that were isolated on immobilised sugars (Gabius, 1991).

In the past, animal lectins have been classified based on the nature of their carbohydrate ligands, the biological processes in which they participate, their sublocalisation and their dependence on divalent cations (Drickamer and Taylor, 1993). However, their classification based on sequence characteristics, specifically their CRDs, is now most widely accepted (Drickamer, 1988, Drickamer and Taylor, 1993, 1998). To date, eight groups of animal lectins have been described. Of these eight groups, four contain lectins that are predominantly intracellular. The intracellular lectins; calnexin family, M-type, L-type and P-type, are located in luminal compartments of secretory pathways. Their functions include trafficking, sorting and targeting of maturing glycoproteins. The four predominantly extracellular lectins include C-type, R-type, siglecs (I-type) and galectins [Figure 1.1]. These four lectin types are typically present in the extracellular matrix, secreted in body fluids, or are localized to the plasma membrane. They mediate a range of functions including cell adhesion, cell signalling, glycoprotein clearance and pathogen recognition. Several additional groups of animal lectins have been proposed, including F-box lectins, ficolins, chitinase-like lectins, F-type lectins and intelectins (Drickamer and Taylor, 1998).

1.1.2.1. Extracellular animal lectins

Among the four predominantly extracellular lectins, the calcium-dependent (C-type) lectins, are involved in sugar recognition in a variety of cell surface and other extracellular lectins (Cummings and McEver, 2009). These lectins are associated with biological processes such as adhesion, endocytosis and complement fixation. The binding of oligosaccharides at the CRD of C-type

lectin involves a Ca^{2+} ion that is essential for the interaction and stabilization of the local conformation of the protein. Members of this family of lectins include lecticans, selectins and collectins, among others.

The R-type lectins, named after the castor bean protein ricin, are the only sugar binding protein modules among animal lectins to be found in prokaryotes (Cummings and Etzler, 2009). Ricin was the first lectin to be identified in this family and is the prototypical protein for this family. It is present in plants, animals and bacteria. The plant lectins usually contain a module that is a potent toxin.

The I-type lectins bind to oligosaccharides through the immunoglobulin superfamily domain. The siglec family of sialic acid – binding cell surface adhesion receptors are the most studied of this family (Crocker *et al.*, 1998). Galectins are globular proteins which are characterised by a galectin–like CRD, which is further discussed in detail below.

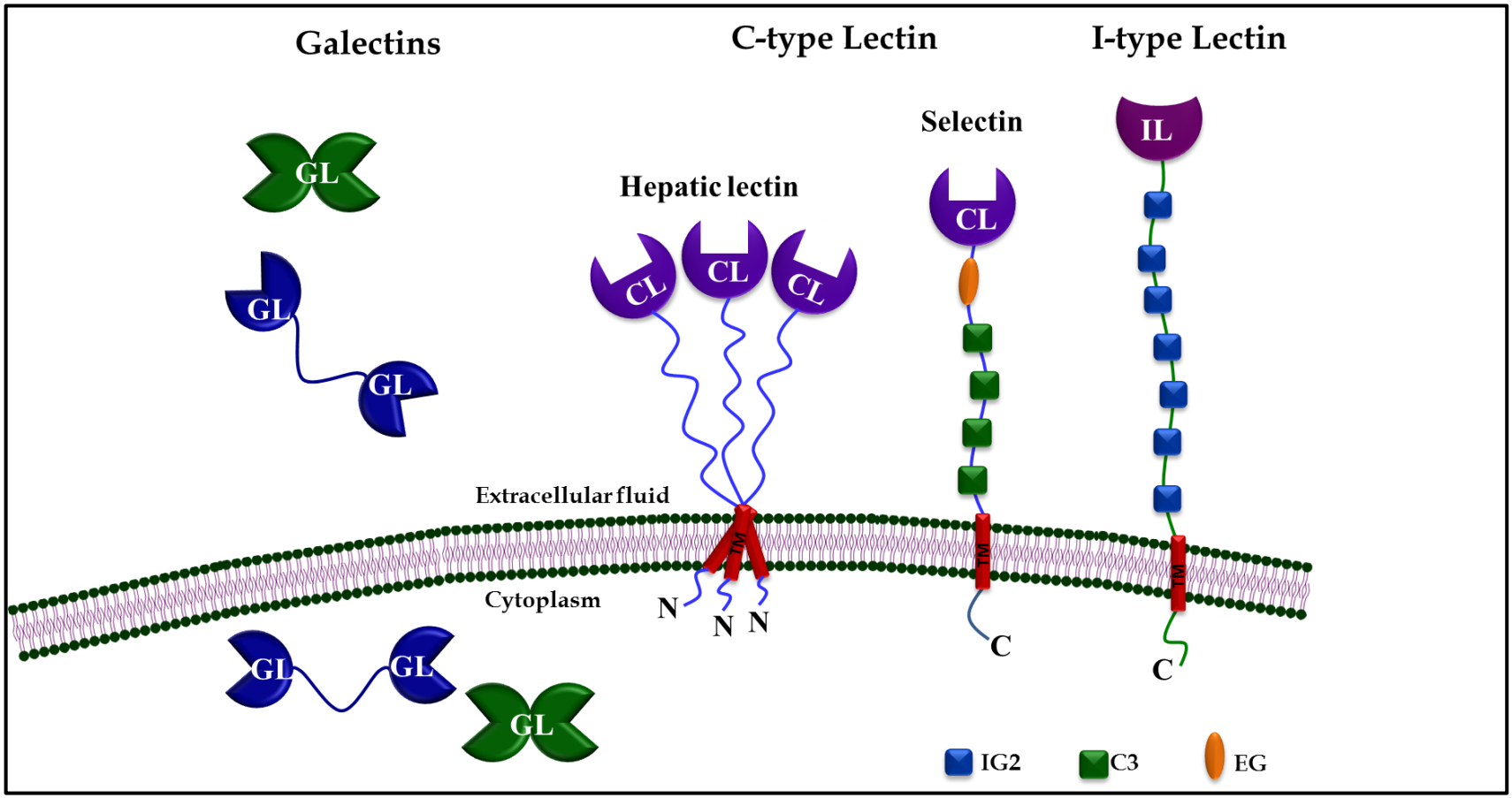


Figure 1. 1: Schematic example of major types of extracellular animal lectins.

Example of domain and topology of some extracellular animal lectins. The defined carbohydrate recognition domains (CRDs) are shown: GL: Galectin CRD, CL: C-type lectins CRD and IL: I-type lectin CRD. Other domains are immunoglobulin C2-set domain (IG2), complement regulatory repeat (C3), transmembrane region (TM) and EGF-like domain (EG). The number of domains underlying the CRD can vary among family members of this family of proteins.

1.1.3. Galectins

Galectins are a family of animal lectins with an evolutionarily conserved CRD, specific for β-galactose containing oligosaccharides, and exhibit high sequence similarity (Barondes *et al.*, 1994b, Cooper, 2002, Danguy *et al.*, 2002, Leffler *et al.*, 2004, Liu *et al.*, 2002, Liu and Rabinovich, 2005, Rabinovich, 1999). They are ubiquitously present in living organisms and are involved in essential cellular events including signal transduction, cell development, differentiation, cell-cell adhesion, cell-matrix interactions, growth regulation and apoptosis (Liu and Rabinovich, 2005, Rabinovich and Toscano, 2009, St-Pierre *et al.*, 2012, Vasta, 2009). The polysaccharide chains present on various cell surfaces and serum proteins play a role in glycan presentation and their distribution has been implicated in various diseases (Gabius, 2000, Rabinovich and Croci, 2012, Reuter and Gabius, 1999, Varki, 1993). The selective recognition of N-glycans and other β-galactosides by members of the galectin family and their involvement in pathological conditions have been extensively studied and highlighted their role in infection, immunity, atherosclerosis and cancer development, through protein-glycan and protein-protein interactions (Delacour *et al.*, 2009, Rabinovich *et al.*, 2002, Rabinovich and Toscano, 2009, Vasta, 2009). The CRD of galectins comprises an α/β sandwich similar to the overall topology of L-type lectins. However, the lack of sequence similarity and the orientation of the sugar-binding sites in these two families suggests that this topological similarity results from convergent evolution. Mammalian galectins lack conventional signal sequences. They reach the cell surface by non-traditional mechanisms before binding to glycoconjugates on the plasma membrane and in the extracellular matrix.

Galectins are synthesised on free polysomes in the cytoplasm and accumulate there prior to secretion. Unlike other animal lectins, they can be found in the nucleus, cytoplasm, outer plasma membrane and the extracellular matrix. Galectins isolated from the cytoplasm of cells bind to β-galactosides which is indicative of their synthesis and functionality in this organelle. Glycan presence is however, not detected in the cytoplasm, which suggests an alternative role

and binding partner of galectins in this compartment (Cummings and Liu, 2009).

Galectins do not contain membrane-anchoring domains or signal peptides required for endoplasmic reticulum (ER)/Golgi-mediated secretion, in their primary structure (Couraud *et al.*, 1989). Moreover, the synthesis of galectins on free ribosomes in the cytoplasm and its subsequent secretion is not blocked by inhibitors such as brefeldin A or monensin which block the ER/Golgi-mediated secretion (Hughes, 1999, Lindstedt *et al.*, 1993, Sato *et al.*, 1993). Therefore, it has been suggested that galectins are secreted through a non-classical pathway in which they first accumulate underneath the plasma membrane of the cell and cause the formation of exosomes [Figure 1.2] (Delacour *et al.*, 2009, Nickel, 2003). Exosomes are vesicles that are formed on the outer surface of the cell in a process termed membrane blebbing (Nickel, 2003). These exosomes are labile structures which then release the protein into the extracellular space.

Galectins are reported to form homogenous lattices at the cell surface through cross-linking that can activate signalling pathways in the control of receptor endocytosis, host-pathogen interactions and activation of immune cells (Brewer, 2002, Vasta, 2009). In human T-cells, the bivalent galectin-1 crosslinks specific glycoprotein counter-receptors on the surface of the cells resulting in apoptosis (Sacchettini *et al.*, 2001).

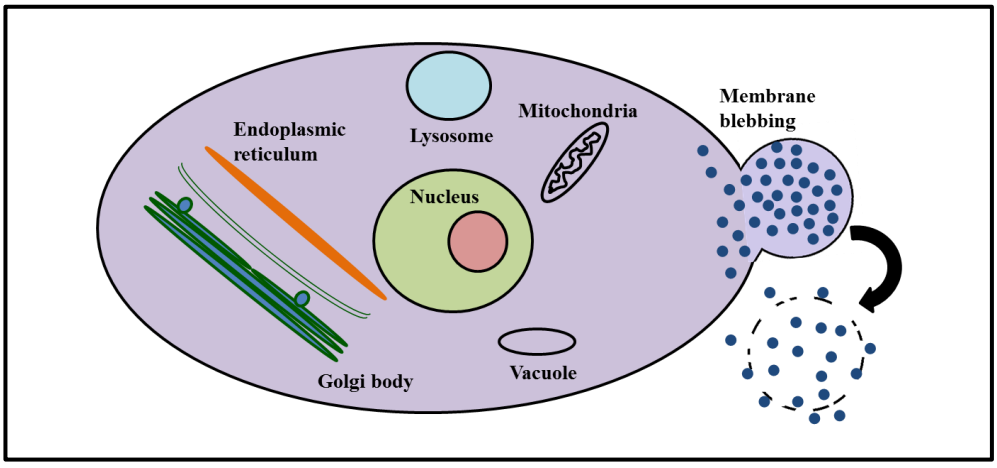


Figure 1. 2: Animal cell undergoing membrane blebbing.
Galectins are secreted through a non-classical pathway involving the formation of exosomes, vesicles that form on the outer surface of the cell.

1.1.3.1 Members of the galectin family

Fifteen members of the galectin family have been identified to date (Gal 1-15) (Hirabayashi and Kasai, 1993). They all have a globular fold consisting of 2 antiparallel β-sheets of 5 or 6 strands and a CRD which forms a concave surface. Based on the structural design of the CRD they are classified into three subgroups: prototype which contain one CRD (Gal -1, -2, -5, -7, -10, -11, -13, -14 and -15), tandem repeat type which contain two CRDs in tandem (Gal -4, -6, -8, -9 and -12) and chimera type which contain one CRD and an additional non-lectin domain (Gal -3) (Hirabayashi and Kasai, 1993, Leffler *et al.*, 2004) [Figure 1.3]. Due to the high level of conservation of the CRD, the disparity in their binding affinity to β-galactosides is characterised by the variation in the surrounding loop regions and β-strands.

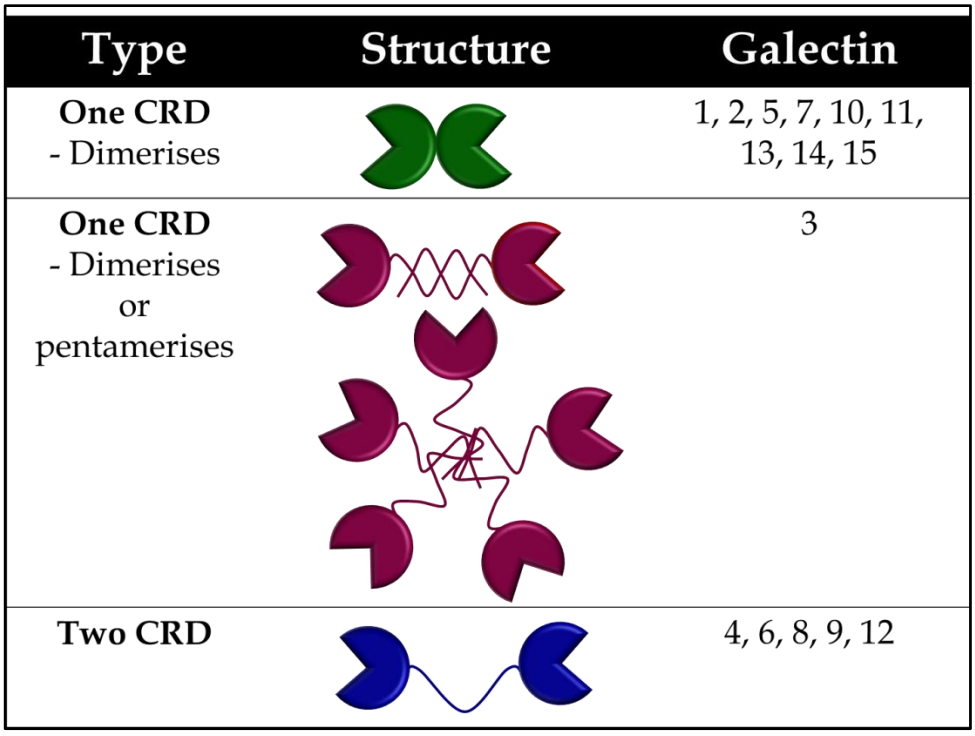


Figure 1. 3: Members of the galectin family.

Most galectins behave as either bivalent or multivalent molecules for ligand binding. The galectins with just one CRD, *i.e.* prototype galectins, tend to dimerise in solution and present as bivalent molecules (Liu and Rabinovich, 2005). The tandem repeat type galectins consist of two CRDs while the chimeric

galectin, *i.e.* galectin-3, oligomerises when it binds to multivalent carbohydrates (Liu and Rabinovich, 2005).

Galectins contribute to cell-cell and cell-matrix interactions. Galectin signalling at the cell surface results in modulation of various cellular functions [Figure 1.4]. Also, in normal cellular physiology, intracellular galectins interact with intracellular ligands and help in the regulation of cellular activities and fundamental processes such as pre-mRNA splicing.

Role of galectins in immune responses

Galectins regulate immune and inflammatory responses. They are expressed by activated T- and B-cells, regulatory T-cells, dendritic cells, mast cells, eosinophils, monocytes/macrophages and neutrophils (Cummings and Liu, 2009). Depending on the inflammatory stimulus, microenvironment and target cells, galectins also promote pro– and anti-inflammatory responses. A short list of the functions and biological roles of galectins in association with their role in immune responses is summarised in the table 1.1.

Role of galectins in apoptosis

Several galectins show the ability to regulate apoptosis. While some galectins can induce apoptosis when added exogenously to cells, others regulate apoptosis through intracellular mechanisms (Liu and Rabinovich, 2005). Galectins affect the process of tumour metastasis through the modulation of cell adhesion and cell migration.

The more explored of these galectins are discussed below.

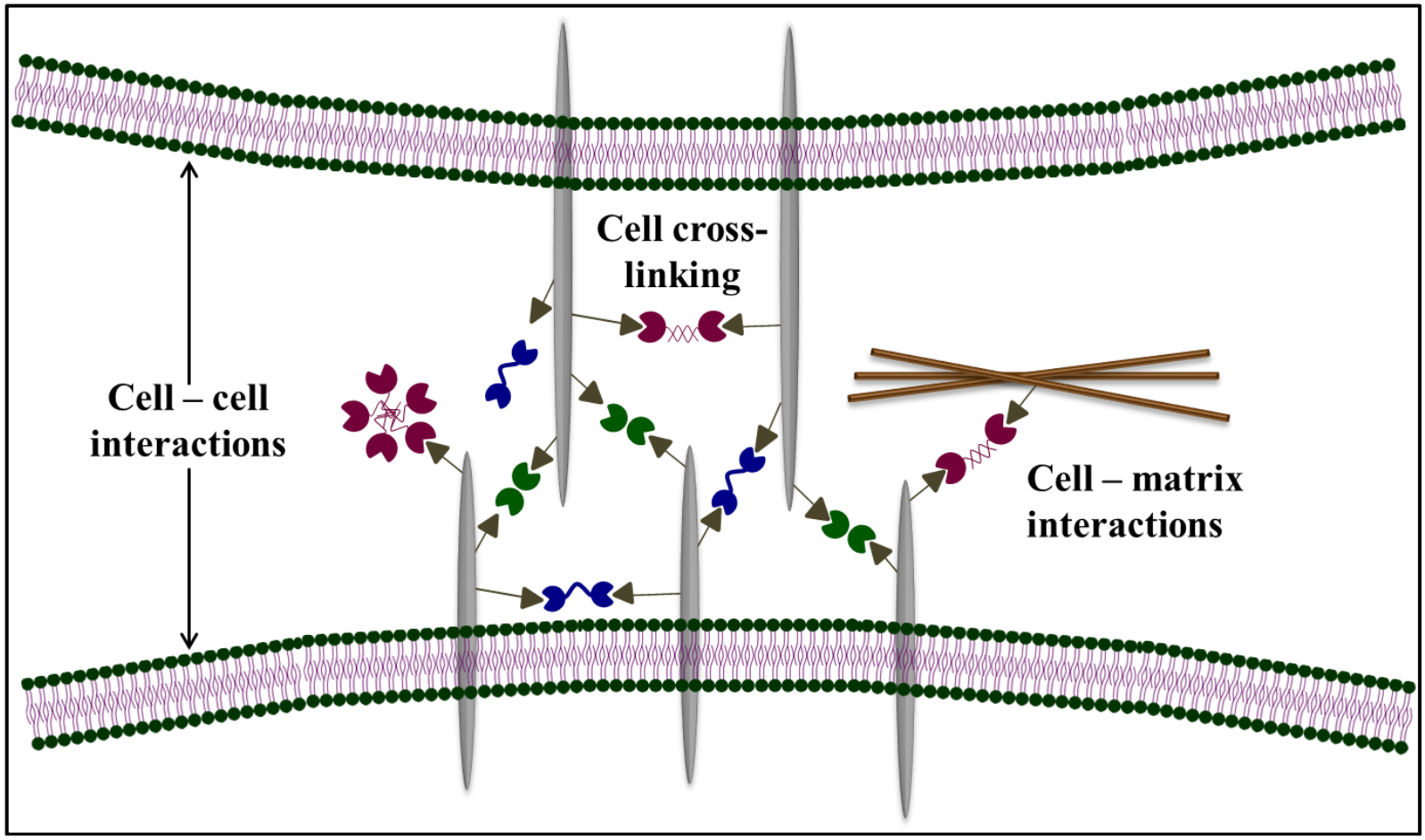


Figure 1. 4: Schematic representation of galectins function

Functional interactions of galectins with cell-surface glycoconjugates and extracellular glycoconjugates can lead to modulation of cell adhesion and cell signalling. Interactions of galectins with intracellular ligands may also contribute to the regulation of intracellular pathways.

Table 1. 1: A list of functions and biological activities of galectins.

Modified from (Cummings and Liu, 2009)

Galectin	Function
Gal-1	reduces TNF-α, IL-1, IL-2, IL-12, IL-2 and IFN-γ increases IL-10 production in T-cells reduces acute inflammatory responses inhibits mast cell degranulation inhibits chemotaxis of neutrophils inhibits extravasation of neutrophils suppression of Th1 and Th17 immune responses activate apoptosis in T-cells induces apoptosis - independent phosphatidyl serine exposure (Ca^{2+} dependent) in neutrophils
Gal-2	induces T-cell apoptosis under some conditions decreases IFN-γ and TNF-α while increasing IL-10 and IL-5 involved in the pathogenesis of atheroma formation induces apoptosis - independent phosphatidyl serine exposure (Ca^{2+} dependent) in neutrophils Binds selectively to β-galactosides of T cells to induce apoptosis
Gal-3	induces release of IL-8 of neutrophils blocks apoptosis of T-cells when overexpressed intracellularly endogenously involved in T-cell viability extracellularly induces apoptosis of T-cells enhances Th2 immune responses inhibits IL-5 production in eosinophils induces mast cell degranulation independent of antigen-mediated IgE stimulation expressed on surface of macrophages exacerbates Th2 immune responses enhances phagocytosis of macrophages enhances respiratory burst of marcophages blocks IL-4 induced survival of activated B-cells favours plasma cell differentiation induces apoptosis intracellularly induces chemotaxis of neutrophils enhances extravasation of neutrophils induces activation of neutrophils enhances leukocyte adhesion to endothelium
Gal-4	induces IL-6 production in T-cells induces apoptosis-independent phosphatidyl serine exposure (Ca^{2+} independent) in neutrophils
Gal-7	intracellular expression induces apoptosis of tumour cells extracellularly can inhibit growth of cells
Gal-8	activates Rac-1 in T cells

Gal-9	activates NADPH- dependent respiratory burst of neutrophils modulates integrin-mediated neutrophil adhesion of neutrophils acts as a danger receptor in defence against intracellular pathogens induces apoptosis in thymocytes and T-cells induces selective loss of CD4+ Th1 cells induces selective loss of CD8+ T cells induces eosinophil chemotaxis, activation, superoxide generation induces moderate degranulation of eosinophil induces maturation of dendritic cells expression in endothelial cells induced by virus infection
Gal-10	highly expressed in eosinophils involved in Treg function
Gal-12	intracellular expression induces apoptosis of tumour cells can cause cell cycle arrest and growth suppression

Galectin-1 (Gal-1) is among the most widely studied member of the galectin family. It is expressed by normal and pathological tissues and is involved in a wide range of biological activities. It is active both intra- and extra-cellularly. The implication of Gal-1 and its ligands in the regulation of immune responses such as T-cell homeostasis and survival, T–cell immune disorders, inflammation, allergies and host-pathogen interaction has been documented (Camby *et al.*, 2006). It also plays an important role in creating immune tolerance in pregnancy and is expressed throughout the menstrual cycle by endometrium stromal cells (Munoz-Suano *et al.*, 2011). Its expression and overexpression in tumours and their surrounding tissues (stroma) is considered a sign of malignancy. Gal-1's other roles include establishment and maintenance of T-cell tolerance and homeostasis *in vivo*, and neuronal cell differentiation and survival in the central and peripheral nervous system. Hence, the involvement of Gal-1 in a myriad of pathological conditions makes it a viable target for novel treatments. Gal-1 has been studied as a potential therapeutic agent in neurological disorders (Chang-Hong *et al.*, 2005, Kadoya *et al.*, 2005) and muscular dystrophy (Cerri *et al.*, 2008, Goldring *et al.*, 2002). The increased expression of Gal-1 and its delivery has been explored for the treatment of inflammation and allied diseases such as arthritis (Rabinovich *et al.*, 1999), nephritis (Tsuchiyama *et al.*, 2000) and colitis (Santucci *et al.*, 2003).

The delivery of anti-Gal-1 compounds in addition to current cytotoxic drugs to down-regulate Gal-1 in cancer related processes is another frontier that is being explored to inhibit metastasis (Ingrassia *et al.*, 2006, Ion *et al.*, 2005, Stowell *et al.*, 2004).

Galectin-2 (Gal-2) is structurally closely related to Gal-1, but its expression and activity are thought to be contained to the gastrointestinal tract (Oka *et al.*, 1999, Sturm *et al.*, 2004). It is an inducer of apoptosis in activated T cells and plays a functional role in T cell homeostasis (Sturm *et al.*, 2004). Gal-2 therefore provides a therapeutic alternative for the treatment of CD8–mediated inflammatory disorders such as contact allergy (Loser *et al.*, 2009). Gal-2 has also been identified as a novel inhibitor of arteriogenesis (van der Laan *et al.*, 2012).

Galectin-3 (Gal-3) contains an unusual tandem repeat of proline and glycine rich stretches of ~120 amino acids, which are fused onto the CRD. It has also been reported that galectin–3 oligomerises *via* its N terminal collagen-like extension after ligand binding (Ahmad *et al.*, 2004, Nieminen *et al.*, 2008). Recent work on galectin-3 showed a different type of self-association upon ligand-binding via its carbohydrate recognition site (Lepur *et al.*, 2012). The formation of galectin oligomers through supermolecular assembly is therefore an essential element of galectin function, which relies on the multivalence of galectins or ligands (Nagae and Yamaguchi, 2014). Gal-3 plays an important role in various biological function and diseases including fibrosis, cardiovascular diseases, inflammatory diseases and cancer. It is commonly up- and down-regulated in different cancers and is implicated in tumour formation and proliferation, apoptosis and B-cell activation (Califice *et al.*, 2004, Ochieng *et al.*, 2004, Stillman *et al.*, 2006). The up-regulation of Gal-3 in liver and renal idiopathic pulmonary fibrosis has been identified. Gal-3 is a proposed prognostic marker in patients with chronic heart failure (Lok *et al.*, 2010). Gal-3 interacts with the poly-N-acetyllactosamine-containing N-glycans on the T-cell receptor and is thus associated with the activation of T-cells (Cummings and Liu, 2009). It is also involved in the inhibition of IL-5 production in immune cells such as human

eosinophils. Moreover, Gal-3 also activates mast cells, neutrophils and monocytes leading to mediator release and production of reactive oxygen species (Cummings and Liu, 2009). Its function as a pro-inflammatory mediator and regulator of inflammatory responses has been widely explored. The modulation of Gal-3 expression in cancer has also been extensively studied.

Galectin-8 (Gal-8) has recently been shown to play a role in cellular defence, against both bacterial cytosolic infection and damage of vacuoles. Thus, Gal-8 acts as a danger receptor in defence against intracellular pathogens (Thurston *et al.*, 2012).

Galectin-9 (Gal-9) plays a role in the regulation of various diseases, including T-cell mediated diseases, such as autoimmunity and asthma. Its involvement in tumour progression and control of antiviral and antibacterial infections has also been reported (Wiersma *et al.*, 2013).

1.1.4. Galectin -7

Galectin -7 (Gal-7) is a prototype galectin which was first identified in 1995 by Madsen *et al.* (1995) in their attempt to identify markers associated with the normal keratinocyte phenotype. In humans, Gal-7 is encoded by the *LGALS7* gene on chromosome 19 (Madsen *et al.*, 1995). Expression of Gal-7 has been shown to be down-regulated in SV40-transformed cells (Madsen *et al.*, 1995) as well as TR146 and SCC13 squamous carcinoma cells (Magnaldo *et al.*, 1995, Magnaldo *et al.*, 1998). A screening, by serial analysis of gene expression, suggested that Gal-7 is coded by 14 of 7202 identified transcripts which were either up- or down-regulated before the onset of p53-induced apoptosis in the colorectal cell line DLD-1 (Polyak *et al.*, 1997). Gal-7, which was known as a marker of epithelial stratification was then referred as PIG1 (p53-induced gene 1) after its role in a model of p53-induced apoptosis (Polyak *et al.*, 1997, Sigal and Rotter, 2000). Subsequently, Gal-7 expression was linked to p53-related induction of epidermal apoptosis by using SCC13 cell transfection using a sense vector and UVB-irradiation as genotoxic stress (Bernerd *et al.*, 1999).

Early reports suggested that Gal-7 existed as a monomer (Madsen *et al.*, 1995). However, the first crystal structure of hGal-7, reported by Leonidas *et al.*, demonstrated that Gal-7 exists as a dimer in crystals (Leonidas *et al.*, 1998). Morris *et al.* (2004) later confirmed the relevance of the dimer complex using ultracentrifugation techniques and a dimer was also observed in the NMR structure reported by Nesmelova *et al.* (2012). Recent analysis by NMR, circular dichroism and molecular dynamic simulations showed that ligand binding to galectin-7 caused long-range effects resulting in the stabilization of the dimer (Ermakova *et al.*, 2013).

The high degree of tissue specificity exhibited by Gal-7 is unique to this member of the galectin family. It is found in various types of epithelial cells such as the mammary myoepithelial cells, the Hassall's corpuscules of the thymus, the urinary system, the epithelial cells of the oesophagus, hair follicles, oral epithelia and the cornea, among others (Demers *et al.*, 2010, Magnaldo *et al.*, 1998, Nio-Kobayashi *et al.*, 2009). The expression of Gal-7 is exclusively exhibited by stratified epithelial cells of the oesophagus, tongue, lip and epidermis (Magnaldo *et al.*, 1998, Sato *et al.*, 2002) and, like other members of the galectin family, is moderately controlled by epigenetic mechanisms such as DNA methylation (Salvatore *et al.*, 2000, St-Pierre *et al.*, 2012). In addition to its regulation by p53 (Polyak *et al.*, 1997), GATA-3 (Demers *et al.*, 2010) and NF-κB (Bemmo *et al.*, 2010, Karin, 2006, St-Pierre *et al.*, 2012) may also regulate Gal-7 expression. However, this exact mechanism of Gal-7 expression and the involvement of the transcription factors in normal cells and/or tumour cells is yet to be determined.

Functions of Gal-7

Gal-7 is expressed with the first onset of epidermal stratification (Magnaldo *et al.*, 1998) and marks the differentiation levels of keratinocyte. It is involved in the migration of epithelial cells and re-epithelisation of corneal and/or epidermal wounds (Cao *et al.*, 2003). The presence of Gal-7 in the areas of cell-cell contact, especially in the upper layers of the epidermis and their

downregulation in transformed keratinocytes, suggests its therapeutic value due to its involvement in physiological cell growth through modulation of cell-cell and cell-matrix interactions.

Gal-7 is distributed uniformly through the epidermis and shows no difference in intensity in the basal and supra-basal layers (Magnaldo *et al.*, 1998). Unlike the complete elimination effect of retinoic acid on standard epidermal markers, the amount of Gal-7 is reduced but not eliminated (Magnaldo *et al.*, 1995), thus distinguishing it from other markers. This exclusivity of Gal-7 is augmented by its unique basal-supra-basal distribution which is not displayed by other markers (Magnaldo *et al.*, 1995, Magnaldo *et al.*, 1998). This suggests that Gal-7 is a keratinocyte cell type marker (Magnaldo *et al.*, 1995).

The possible role of Gal-7 in cell-mucus contact and growth of the ovarian follicles has been suggested due to its presence in the trachea and the ovary (Sato *et al.*, 2002). Gal-7 modulates cell proliferation and cell interactions irrespective of keratinisation or regional specialisation and is thus important for both the formation and maintenance of all stratified epithelia (Timmons *et al.*, 1999). Recently, Gal-7 and actin were detected in the water-soluble fractions prepared from skin lesions of patients with primary and secondary localized cutaneous amyloidosis and have been considered as candidates with therapeutic value (Miura *et al.*, 2013).

Gal-7, in addition to Gal-3, but unlike Gal-1, has been shown to stimulate the re-epithelialization of corneal wounds (Cao *et al.*, 2002a, Cao *et al.*, 2003, Cao *et al.*, 2002b). The extent of acceleration of re-epithelialization of wounds was reported to be greater than that observed in most of the published studies using growth factors (Cao *et al.*, 2002a). Additionally, Gal-7 and Gal-3 have been shown to induce cell mitosis in epithelial cells (Cao *et al.*, 2002a, Cao *et al.*, 2003). Thus, these studies provide a platform for the development of novel therapeutic tools for the treatment of non-healing wounds.

Gal-7 expression in women has been studied extensively in recent years. It is produced by the premenstrual and menstrual endometrium. Here, it facilitates

post-menstrual endometrial re-epithelialization by accumulating in the menstrual fluid and acts as a paracrine factor (Evans *et al.*, 2014). An abnormal elevation of Gal-7 levels in the endometrium of women with a history of miscarriage is often observed. Studies suggest that Gal-7 facilitates the adhesion of the embryo to the endometrium and elevated levels of Gal-7 result in abnormal adhesion, which can lead to miscarriage (Menkhorst *et al.*, 2014b).

Additionally, due to the elevated levels of Gal-7 levels in the serum of pregnant women who subsequently developed pre-eclampsia, compared to women with healthy pregnancies, it is considered as a prognostic serum biomarker for this condition (Menkhorst *et al.*, 2014a).

Role in cancer

The modulatory effect of galectins on apoptosis has been widely studied. This modulation of apoptosis is either through the regulation of apoptosis by binding to cell surface glycoproteins or through the interactions with relevant intracellular proteins (Liu and Rabinovich, 2005). While Gal-3 has anti-apoptotic activity, Gal-1,-2, -7 and -9 show pro-apoptotic activity (Saussez and Kiss, 2006). *De novo* expression of Gal-7 by p53 is associated with apoptosis, thereby inhibiting cancer progression. Gal-7 has a pro-apoptotic intracellular function that plays a role upstream of the c-Jun N-terminal kinases (JNK) activation pathway and cytochrome c release (Kuwabara *et al.*, 2002). Gal-7 can also inhibit cell proliferation without its pro-apoptotic role. The inhibition of proliferation in neuroblastoma cells by mediation of Gal-7 has been attained without any signs of apoptosis (Kopitz *et al.*, 2003). The improved outcome of cervical cancer patients after radiotherapy has been shown to be associated with high levels of Gal-7 (Tsai *et al.*, 2013). Gal-7 is also significantly down-regulated in gastric cancer as the gene is epigenetically modified by DNA methylation (Kim *et al.*, 2013c). Thus, the role of Gal-7 as a tumour suppressor with potential therapeutic activity supports its expression being induced by p53, an important tumour antigen.

Over the recent years, in addition to the tumour suppressive role of Gal-7, its tumour progressive effects have also come to light. Gal-7 is known to interfere with Smad3, a key regulatory protein in the transforming growth factor – β (TGF-β) signalling pathway which is known to exert tumour suppressive and promoting effects (Inagaki *et al.*, 2008). Gal-7 has also shown to aid tumour growth by not inducing apoptosis in certain cases. Studies have shown that Gal-7 expression is induced in breast cancer by mutant p53 (Campion *et al.*, 2013). Gal-7 is overexpressed in mammary tumours compared to normal mammary tissues, when subjected to chemical induction of mammary carcinoma (Lu *et al.*, 1997). Increased levels of Gal-7 have been observed in mature B-cell lymphoid neoplasms but not in normal B-lymphocytes (Demers *et al.*, 2007). Gal-7 has also been shown to be a negative growth regulator for neuroblastoma cells (Kopitz *et al.*, 2003). High Gal-7 levels have been associated with poor survival outcome in patients with epithelial ovarian cancer, in addition to being functionally involved in cell proliferation (Kim *et al.*, 2013a). Gal-7 moderates the aggressive behaviour of lymphoma cells by modulating the expression of a metastic gene, matrix metallopeptidase (MMP-9) (Demers *et al.*, 2005) and increases metastasis. Induction of MMP-9 by galectins modulates the microenvironment by promoting integrin mediated cell adhesion and migration of tumours cells, that may be exploited by cancer cells. In such cases, Gal-7 expression may be induced by nuclear factor kappa-light-chain-enhancer of activated B-cells (NF-κB), a gene expressed in highly aggressive cancer cells and a positive regulator of MMP-9 (St-Pierre *et al.*, 2012).

The dual role of hGal-7 [Figure 1.5] in the modulation of tumour and cancer cell migration provides an excellent therapeutic opportunity for understanding and controlling tumour growth and metastasis. However, more studies are needed to further delineate the precise role of hGal-7 in cancer. Identifying and understanding specific environment, secondary structure, oligomerisation or supramolecular formation and binding partners of hGal-7 at each stage of cancer progression will help precise modulation of hGal-7 and in turn aid rational design of novel therapeutics against cancer.

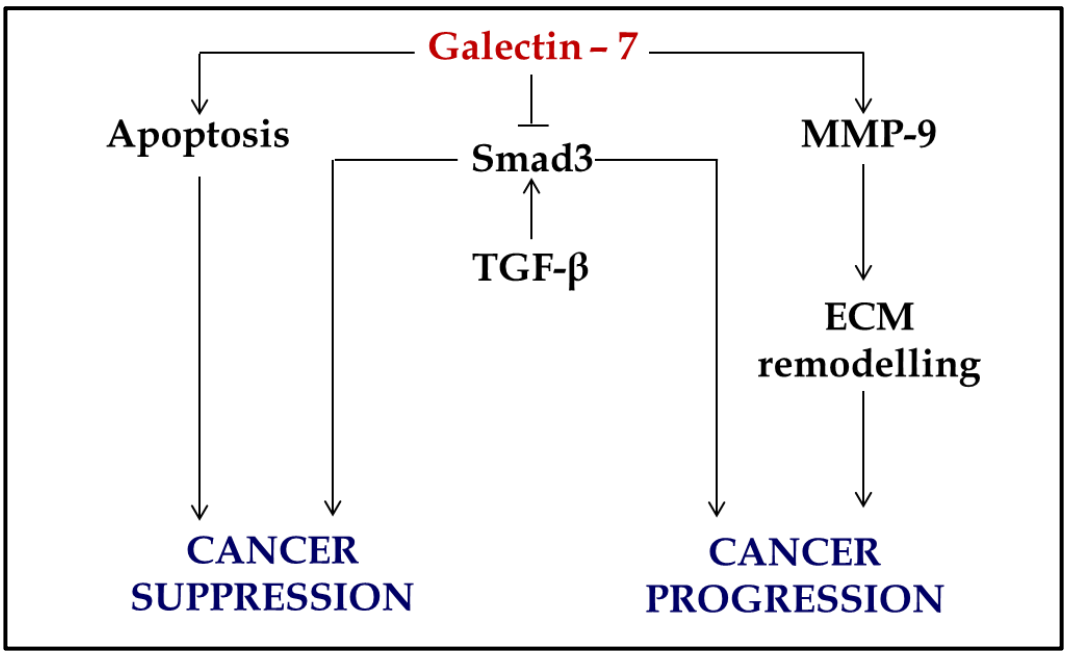


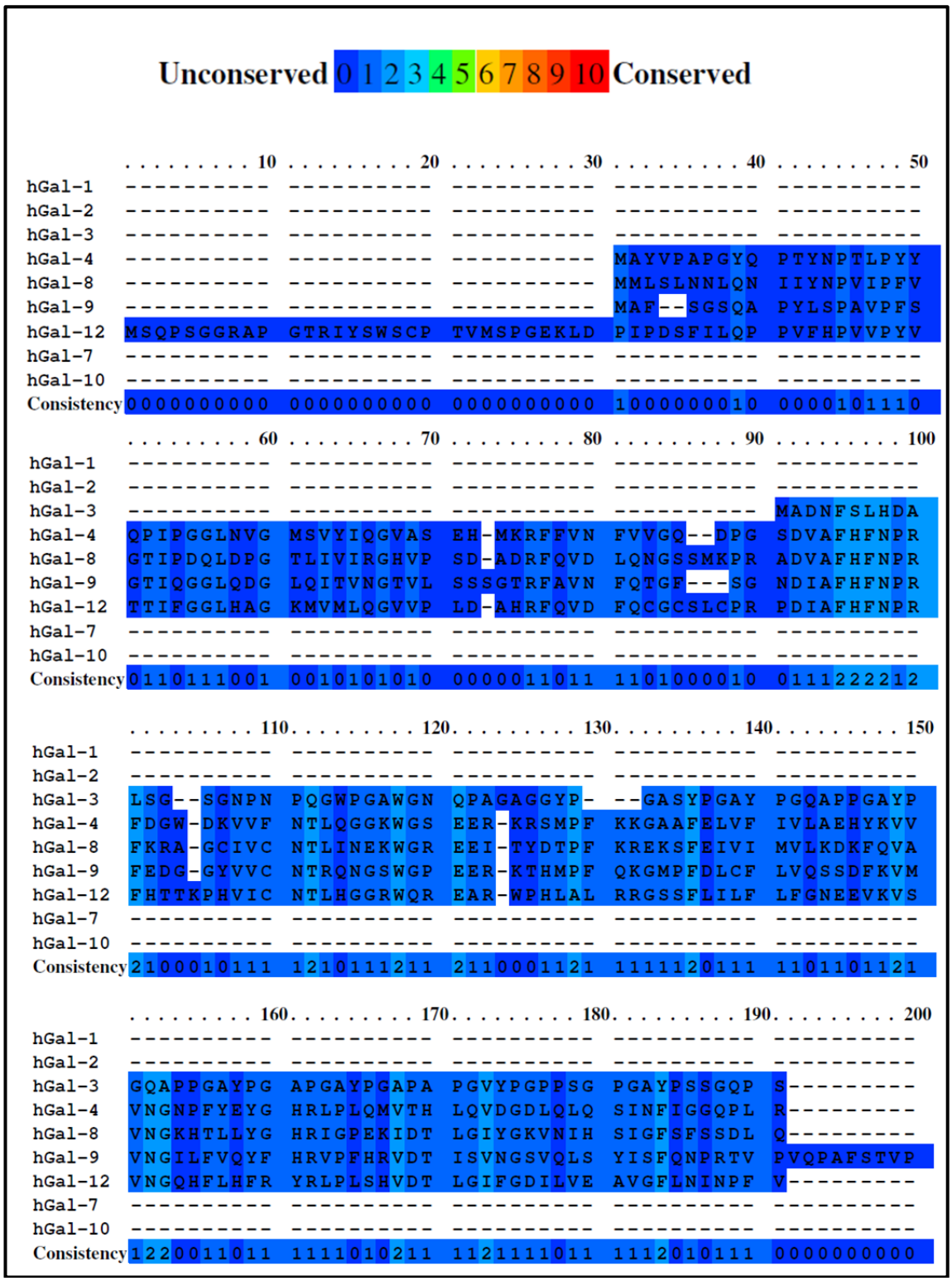
Figure 1. 5: Dual role of hGal-7 in cancer.

hGal-7 plays a role in tumour suppression as well as its progression. Its expression by p53 is associated with apoptosis and inhibition of tumour growth. hGal-7may also modulate cancer progression by interfering with Smad3, an important regulatory protein of TGF-β pathway which is involved in both, tumour suppression and tumour progression. Induction of hGal-7 expression by NF-κB can in turn induce MMP-9, which is involved in modulation of the extracellular matrix (ECM) environment, that cancer cells may exploit to their advantage. Modified from (St-Pierre *et al.*, 2012)

1.2. Structural features of galectins

The structures of several galectins have been elucidated at the atomic level. The protein data bank (PDB) includes approximately 170 X-ray and NMR structures of full or specified individual domains of galectins in their native form or in complex with various ligands. These galectin structures are from a wide range of organisms which include *Homo sapiens* (human), *Agrocybe aegerita* (fungus), *Mus musculus* (house mouse), *Coprinopsis cinerea* (fungus), *Porcine adenovirus*, and several others. Significant similarities between the tertiary structures of lectins from various species have been noticed in spite of the lack of primary sequence similarity. The structures of human galectins that have been reported include crystal or NMR structures for native and/or complexes with ligands for hGal - 1, 3, 4, 7, 8, 9.

Galectins are known to be involved in signal mediation, cell-cell and cell-matrix interactions which may be brought about by the cross-linking of N-acetyllactosamine-containing structures found at cell surfaces and in the extracellular surfaces. Generally, the canonical CRD of galectins consists of nearly 130 amino acids of which only a few residues directly interact with the ligand. A PRALINE (Bawono and Heringa, 2014) sequence and ConSurf (Celniker *et al.*, 2013) analysis of these residues reveals conservation primarily of the inward facing hydrophobic residues in β-strands in the β-sandwich of the galectin fold [Figures 1.6 and 1.7]. Sequences of the galectins were obtained from the Universal Protein Resource (UniProt, 2014).


Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved
hGal-1
hGal-2
hGal-3
hGal-4
hGal-8
hGal-9
hGal-12
hGal-7
hGal-10
Consistency

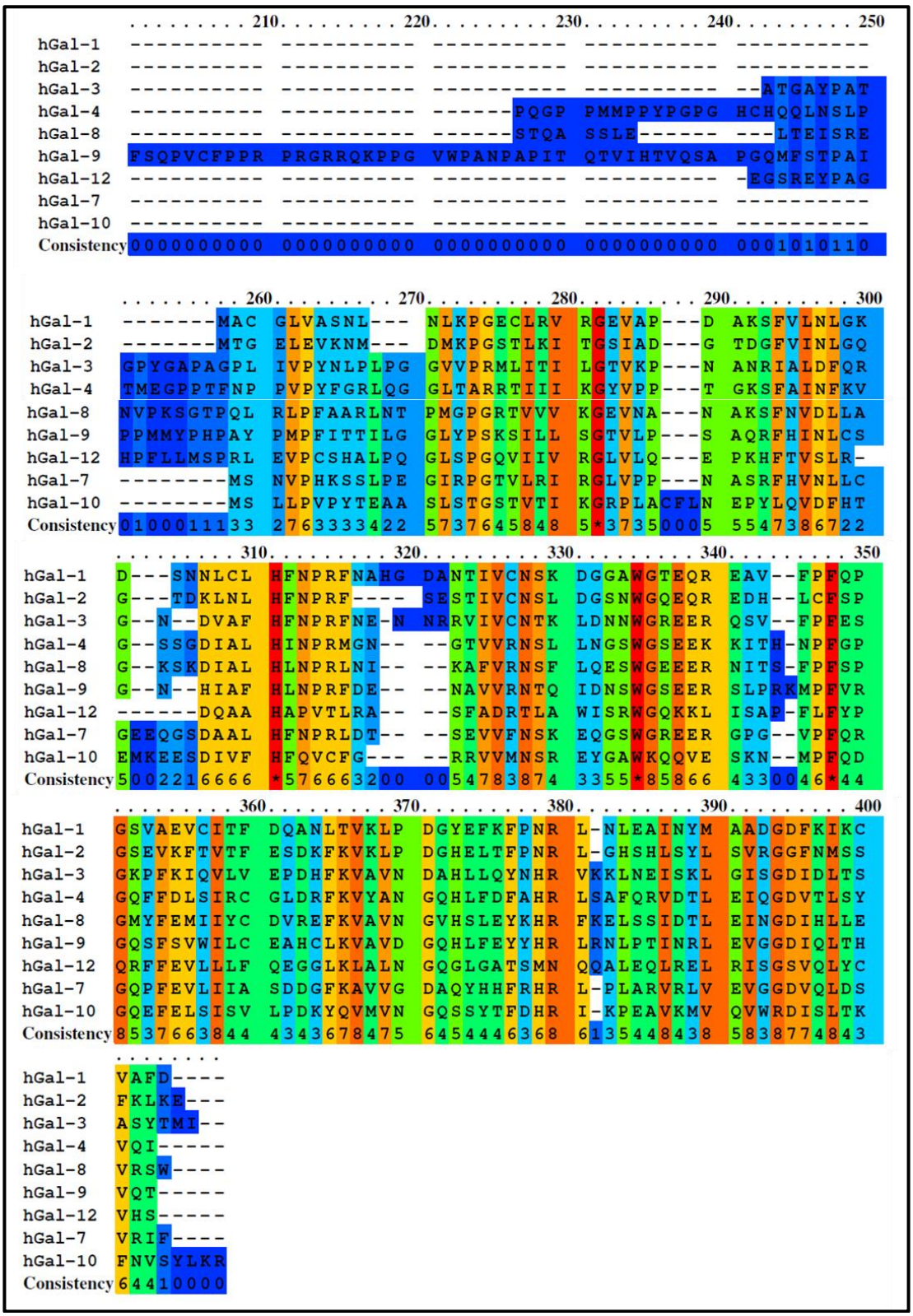


Figure 1. 6: PRALINE sequence analysis of hGal-1, hGal-2, hGal-3, hGal-4, hGal-7, hGal-8, hGal-9, hGal-10 and hGal-12.

The analysis elucidates the high degree of conservation of the CRD composed of residues ~305 – 400. The residues are coloured by their conservation grades using the colour coding bar, with blue-through-red indicating unconserved-through-conserved residues.

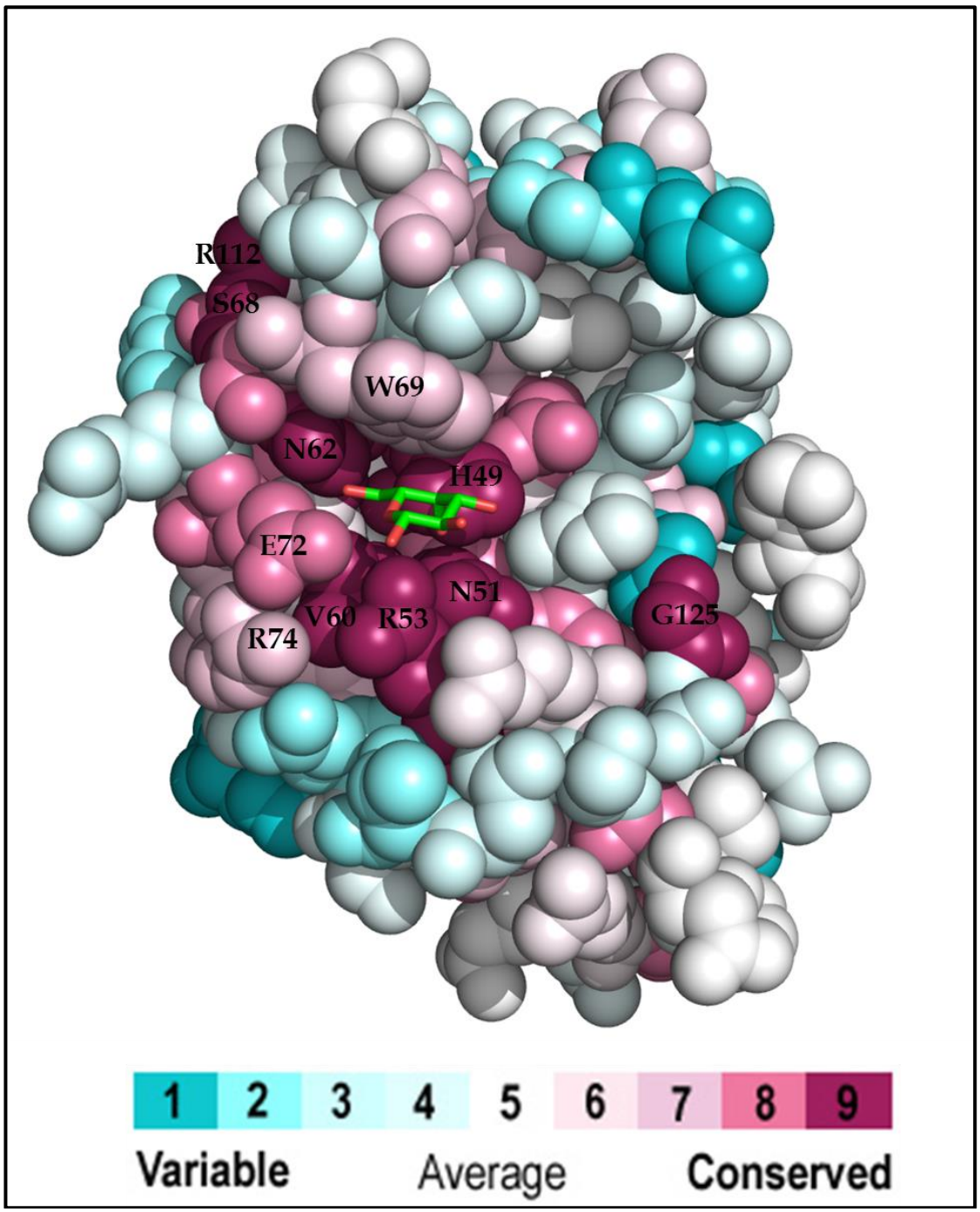


Figure 1. 7: A ConSurf analysis of hGal-7 in complex with galactose

(PDB code: 2GAL). Galactose is shown as sticks in green bound at the CRD. The amino-acids are coloured by their conservation grades using the colour-coding bar, with turquoise-through-maroon indicating variable-through-conserved residues.

The galectin CRD is composed of five and six stranded antiparallel β-sheets arranged in a β-sandwich or a 'jelly-roll' topology that completely lacks α-helices [Figure 1.8]. The monomeric subunits of dimeric prototype galectins: hGal-1, hGal-2 and hGal-7 are related by a two-fold rotational axis perpendicular to the plane of the β-sheets. The N- and the C- termini of each monomer are generally positioned at the dimeric interface. The CRDs in the dimer are located at opposite ends of the dimer and constitute a long and negatively charged cleft in the cavity. The compact arrangement of the CRD elucidates the high degree of conservation of this domain .

A structural alignment of the galectins using GESMAT (Krissinel, 2012) revealed a high degree of conservation of the secondary structure [Figure 1.9]. The corresponding residues to H49, R53, N62, W69, E72 and R74 that interact with the ligand at the CRD of hGal-7 are observed at equivalent positions in the other human galectin structures. Table 1.2 lists the corresponding residues of galectins involved in ligand binding. Both, secondary structural and primary sequential conservation of these residues are observed.

Table 1. 2: Residues of hGal-7 that interact with the ligand with the corresponding residues in other human galectin structures.

Galectin	PDB code	H	R	N	W	E	R
1	**3W59**	44	48	61	68	71	73
3	**3ZSM**	158	162	174	181	184	186
7	**3ZXF**	49	53	62	69	72	74
8	**2YV8**	235	239	248	255	258	260
9	**3NV1**	72	76	86	93	96	-

Gal-1 (PDB code: 3W59) (Saburi *et al.*, unpublished results), Gal-3 (PDB code: 3ZSM) (Saraboji *et al.*, 2012), Gal-7 (PDB code: 3ZXF) (Masuyer *et al.*, 2012), Gal-8 (PDB code: 2YV8) (Kishishita *et al.*, unpublished) and Gal-9 (PDB code: 3NV1) (Yoshida *et al.*, 2010).

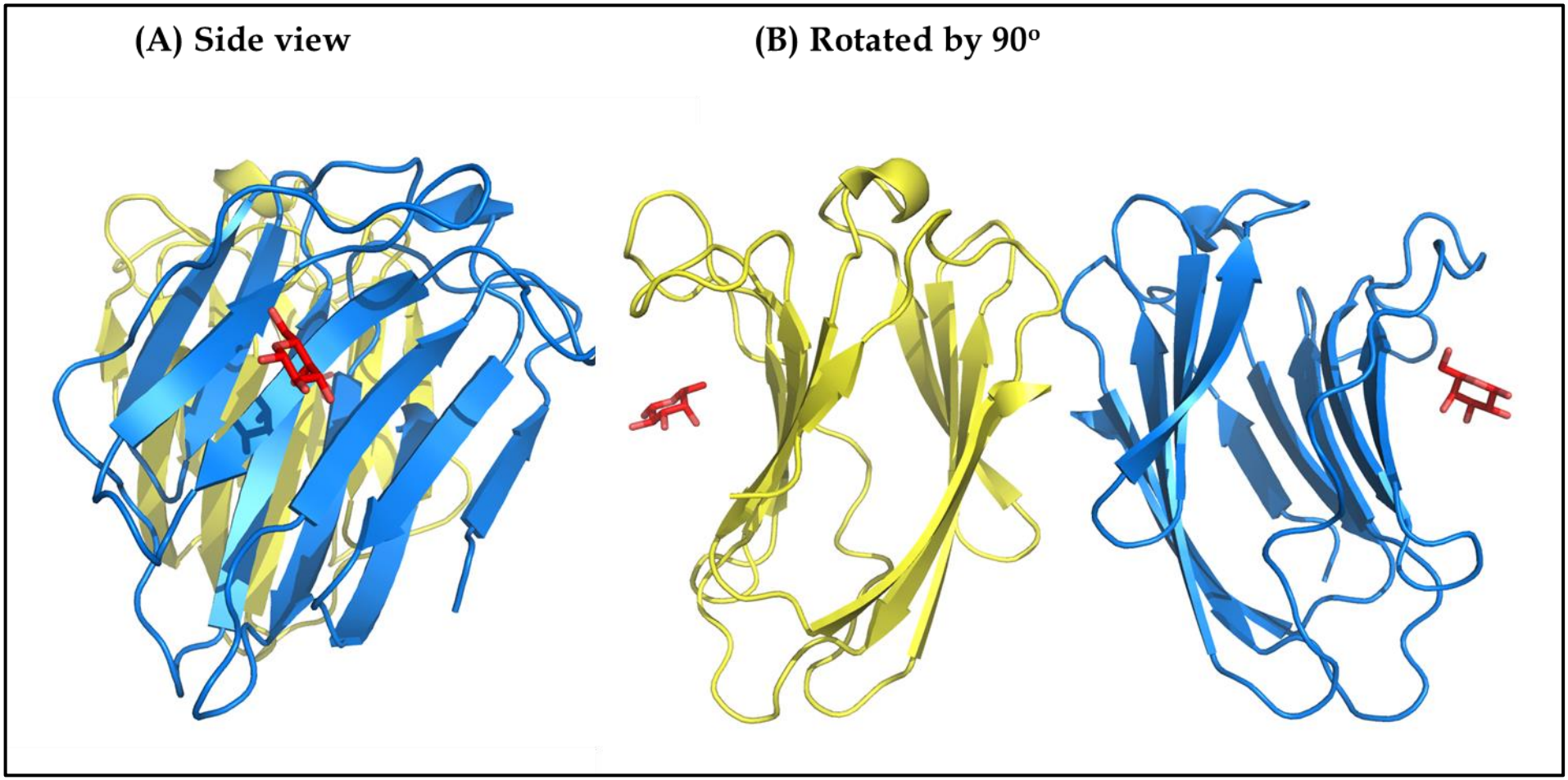


Figure 1. 8: Ribbon diagram of the crystal structure of hGal-7 in complex with galactose

(PDB code: 2GAL). The homodimer is shown. Each monomer is coloured differently and orthogonal views have also been presented.

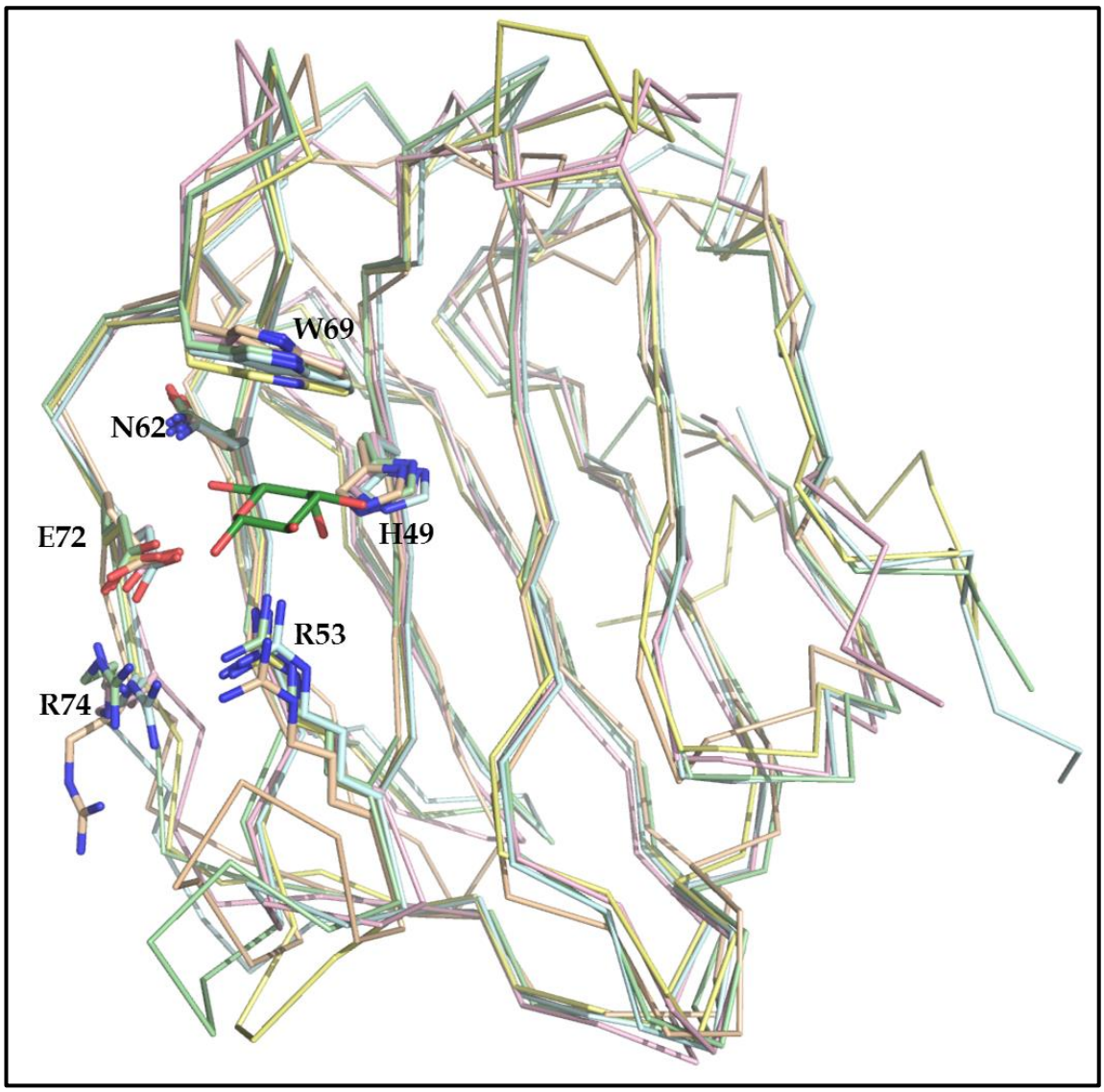


Figure 1. 9: Superposition of hGal-1, -3, -8 and -9 with respect to hGal-7.

The residues involved in the binding of the ligand (galactose in green) at the CRD are shown as sticks and labelled as per hGal-7 residue numbering. A high degree of secondary structural features among galectins are observed including the important residues involved in ligand binding at the CRD. The PDB codes for hGal-1 (wheat), hGal-3 (pale green), hGal-7 (light pink), hGal-8 (light yellow) and hGal-9 (light blue) are 3W59, 32SM, 3ZXF, 2YV8 and 3NV1, respectively.

1.3. Carbohydrates

Carbohydrates are one of the four major classes of biomolecules along with proteins, nucleic acids and lipids. They play a vital role in mediating interactions among cells and between cells and other elements in normal cell physiology. For this purpose, they are either linked to proteins or lipids. Glycoproteins are proteins derivatised with carbohydrates. They are components of the cell membranes. Their structural diversity and location within proteins enables carbohydrates to act as mediators in a wide range of cellular interactions. The extensive structural polymorphism of long chain carbohydrates is due to variations in its monosaccharide composition, carbohydrate sequence, branching, linkage position and linkage anomericity.

Carbohydrates are carbon compounds with an empirical formula $C_m(H_2O)_n$, (where m may or may not be equal to n) forming a "carbon-hydrate" in most cases. Monosaccharides form the building blocks of complex carbohydrates. They are aldoses or ketoses with multiple hydroxyl groups. Two to ten monosaccharides linked by glycosidic bonds make up an oligosaccharide while hundreds of monosaccharide units form polysaccharides. The abundance of hydroxyl groups aid their interaction with the aqueous environment and facilitate in hydrogen bonding between chains. Mostcarbohydrates contain at least one chiral (asymmetrical) carbon and are therefore optically active. Generally speaking, only the D-enantiomer of a carbohydrate is physiologically relevant, albeit with a few exceptions. The mirror image conformation, called an L-enantiomer, The enantiomers of glyceraldehyde are shown below [Figure 1.10].

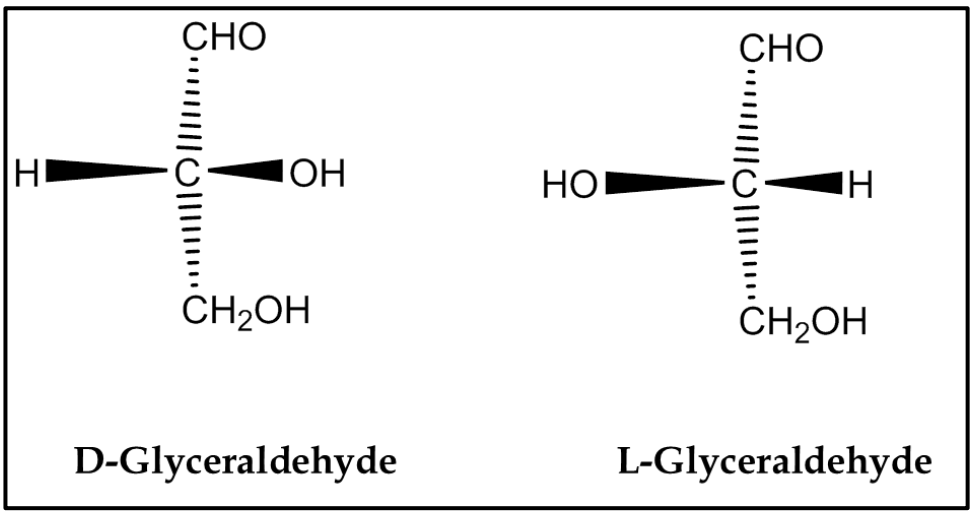


Figure 1. 10: Fisher projections of the enantiomers of glyceraldehyde.

Monosaccharides

Monosaccharides commonly found in humans are classified based on the number of carbon atoms they contain in their backbone structure [Table 1.3]. The major carbohydrates contain four to six carbon atoms.

Intramolecular hemiacetals and hemiketals are formed when the aldehyde or ketone moieties of the carbohydrate react with alcohol groups present in neighbouring carbons. This results in the formation of five or six-membered rings which are called furanose or pyranose forms, respectively.

The opening and re-closing of the rings results in freedom of rotation about the carbon bearing the reactive carbonyl. This rotation yields two distinct configurations (called anomers) of the hemiacetals and hemiketals: α- and β-anomers. The carbon about which this rotation occurs is referred to as the anomeric carbon and the process of spontaneous change between these two configurations is termed mutarotation. Figure 1.11 shows the cyclic and chair projection of β-D-glucose.

Table 1. 3: Carbohydrate classification

No. of Carbons	Category	Examples
3	Triose	Glyceraldehyde
4	Tetrose	Erythrose
5	Pentose	Ribose, Ribulose
6	Hexose	Glucose, Galactose
7	Heptose	Sedoheptulose
9	Nonose	Neuraminic acid (a.k.a. sialic acid)

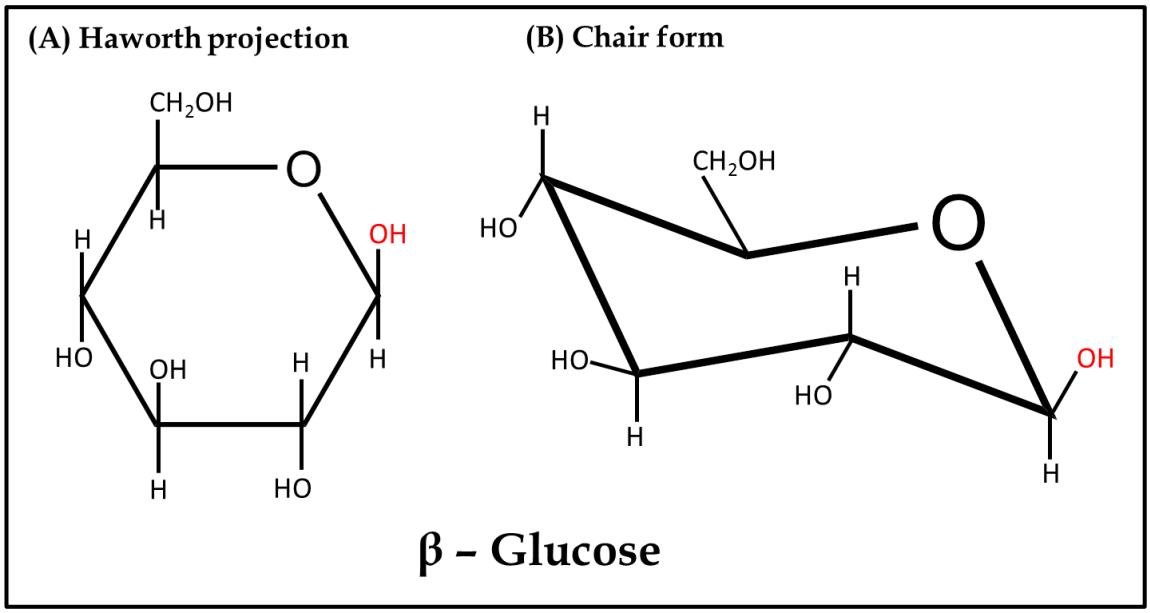


Figure 1. 11: Cyclic and chair conformation of β-D-glucose.

Cyclic projection given for simplicity, while the chair conformation is an example of a more biologically relevant conformation.

Disaccharides

The linkage of two monosaccharides through a glycosidic bond results in a disaccharide. Glycosidic bonds are covalent bonds between the anomeric hydroxyl of a cyclic sugar and a hydroxyl of a second sugar or another alcohol containing compound. Lactose [Figure 1.12], found in the milk of mammals, is a disaccharide is made up of galactose and glucose linked via a β-(1,4) glycosidic bond.

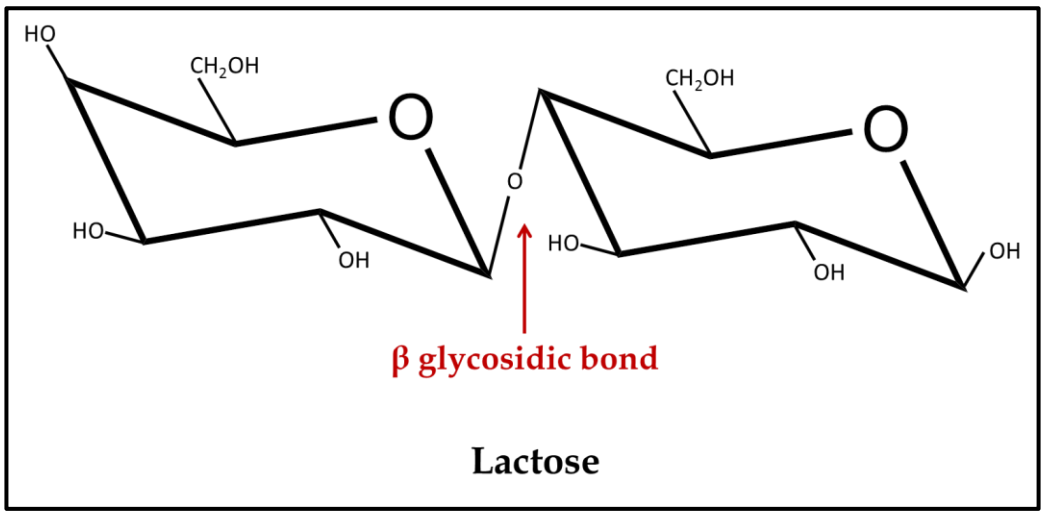


Figure 1. 12: Structure of lactose.

Oligosaccharides

A saccharide polymer containing a small number of monosaccharides (usually two to ten) are called oligosaccharides. Each monosaccharide is linked to the adjoining sugar via glycosidic bonds. Oligosaccharides are commonly found on the plasma membrane of animal cells and play an important role in cell-cell recognition.

Oligosaccharides are components of glycoproteins and glycolipids. They are used as chemical markers and for cell recognition. The blood type specificity is dictated by the oligosaccharide glycolipid forming the cell membrane of the red blood cells. Blood type A and type B have different pentasaccharides while type AB have both and type O has a tetrasaccharide.

As carbohydrates are recognised by various proteins, carbohydrate based tools such as glycodendrimers used in this thesis, are used in a diverse range of explorative studies to decipher the mechanistic role of proteins.

1.3.1. Glycodendrimers and galectins

Dendrimers are synthetic, monodisperse macromolecules consisting of three distinguishing structural features: the core, the branches or arms of the molecule and the termini which can all harbour a variety of chemical functions (Newkome *et al.*, 2001, Villalonga-Barber *et al.*, 2008, Vögtle *et al.*, 2009). The polyvalent nature of dendrimers and their flexibility make them unique nanodevices for biochemical applications (Boas *et al.*, 2006, Gardikis *et al.*, 2012).

Glycodendrimers constitute a class of dendrimers that incorporate carbohydrate into their structures (Roy *et al.*, 1993) and have potential biomedical applications such as, for example, anti-adhesion drugs (Touaibia and Roy, 2007), drug delivery systems (Chabre and Roy, 2012), functional antigens and antitumor vaccines (Shiao and Roy, 2012, Teo *et al.*, 2012), biosensors (Kikkeri *et al.*, 2009, Kikkeri *et al.*, 2010a) and gold nanoparticle conjugates (Gorityala *et al.*, 2012, Lyu *et al.*, 2008, Reynolds *et al.*, 2006). Inhibition of the cholera toxin through recognition of the carbohydrate group has also been reported (Branderhorst *et al.*, 2007), demonstrating the potential of these molecules as direct interacting partner for a protein target of choice.

The ability of amphiphilic dendrimers designed with either a D-mannose, D-galactose or D-lactose termini, named Janus glycodendrimers, to self-assemble in water has been demonstrated (Percec *et al.*, 2013). The results revealed the potential use of glycodendrimers as mimics of biological membranes with programmable glycan ligand presentations as supramolecular lectin blockers, vaccines and targeted delivery devices. The binding affinity of glycodendrimers with lectins is enhanced compared with the corresponding monosaccharides and this has been ascribed to a "*cluster glycoside effect*" (Dimick *et al.*, 1999, Mouline *et al.*, 2014, Varki, 1993), and led to the synthesis of many dendrimeric glycoconjugate compounds (Kikkeri *et al.*, 2010b, Roy and Kim, 1999, Wolfenden and Cloninger, 2005). The polyvalent nature of dendrimers may help in selective modulation of galectins as diverse substrate specificity and binding affinity is observed among its various members (Hirabayashi *et al.*, 2002).

The cross-linking of galectin-1, -3 and -7 by dendrimeric glycoconjugates in solution has been reported (Andre *et al.*, 1999, Andre *et al.*, 2001, Michel *et al.*, 2014, Percec *et al.*, 2013). Glycodendrimers have been studied as inhibitors of binding of mammalian galectins (galectin-1, -3 and -7) to glycoproteins, lactose clusters and cell-surface glycoconjugates with relevance to inflammation and tumour embolisation (Andre *et al.*, 2001). Michel *et al.*, reported that glycodendrimers could mediate the interaction between galectin-3 and mucin-1 which is overexpressed and highly glycosylated in cancer cells (Michel *et al.*, 2014). The tunable multivalent lactose-functionalised dendrimers used in this study was shown to modulate the level of cellular aggregation in lung carcinoma, prostate cancer and fibrosarcoma cell lines.

The details of the dendrimers used in this thesis are given in section 4.2.

1.4. Aims and objectives

Human Galectin-7 (hGal-7) is a prototype galectin which dimerises in solution. The X-ray crystal structure of hGal-7, in which its dimerisation was first observed, was first elucidated by Leonidas *et al.* (1998). Ultracentrifugation techniques and NMR analysis later confirmed the dimerisation of hGal-7 in solution (Morris S *et al.*, 2004, Nesmelova *et al.*, 2012). Its involvement in cell – cell and cell – matrix interactions has highlighted their importance in signal transduction and other intracellular processes. The natural substrates of hGal-7 are yet to be identified. It does however show a high affinity for galactose *in vitro*. With its inherent self-assembling capability, hGal-7 represents an ideal prototype lectin for studying the formation of ligand-associated supramolecular assemblies.

Monovalent binding between galectins and carbohydrates is of relative low affinity. However, polyvalent crosslinking of glycoconjugates leads to the formation of galectin-carbohydrate lattices of high affinity at the cell surface (Brewer, 2002, Rabinovich *et al.*, 2007) as shown in Figure 1.13.

The study of galectin and carbohydrate based dendrimers presents a novel path to design new tools to investigate the recognition of carbohydrates by galectins and their supramolecular formations. Understanding the structural composition of these complexes and the binding affinity of hGal-7 to glycodendrons will provide a platform for the design of molecular markers for cellular interactions and as therapeutic targets. X-ray crystallography is considered the gold standard to decipher and understand these interactions at the atomic level. Molecular details unearthed using this technique can then be applied for the rational design of synthetic, highly-specific molecules to target lectins.

In order to perform detailed biochemical and structural analyses, the expression and purification of recombinant hGal-7 was initially optimised. To investigate the supramolecular formation of hGal-7 through carbohydrate recognition, synthetic multivalent galactose based dendrons of varying size and

composition were employed. The purified protein in complex with a range of ligands, was prepared for investigation by X-ray crystallography as a means to understand the recognition of dendrons and the resulting lattice formations. Additionally, isothermal titration calorimetry studies were performed to study the binding affinity of glycodendrons to hGal-7. This led to preliminary binding affinity studies which may offer a basis for future investigations. Dynamic light scattering experiments were also performed to assess the supramolecular assembly formation in solution.

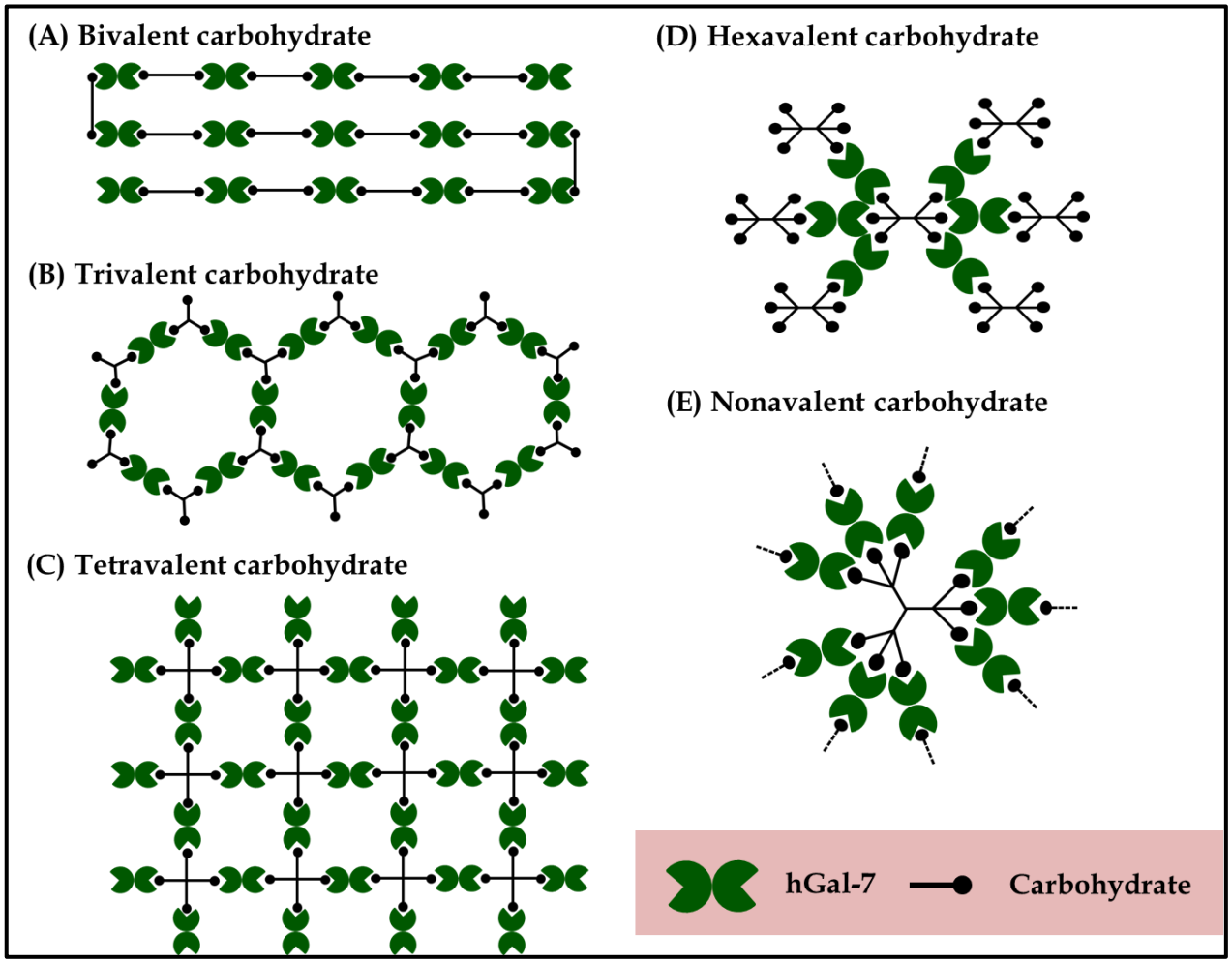


Figure 1. 13: Lattice formation of hGal-7 on interacting with multivalent carbohydrate based ligands.

Chapter 2. INTRODUCTION TO SIRTUINS

2.1. Overview

Sirtuins have been widely explored as elixirs of life. Their involvement in numerous cellular functions in normal cell physiology and pathological conditions make them a fascinating subject of scientific study. They form part of a feedback mechanism that enhances cell survival during times of stress, for example during calorie restriction, and they serve as guardians of the cell. Sirtuins have been extensively explored to understand their role in ageing and ageing related diseases. This chapter provides a brief introduction to the present understanding of sirtuins.

2.1.1. Discovery of the Sir2 family

Nearly eighty years ago, Clive McCay *et al.* first reported that restriction of calories without malnutrition prolongs lifespan in rats compared with *ad libitum* feeding (McCay *et al.*, 1935). Despite the advances in science since then, the molecular mechanism that drives lifespan has been elusive. In addition to rats, many other model organisms have been used to study the molecular mechanism of ageing.

One such model is yeast, a simple eukaryotic organism that shows a high degree of conservation of numerous essential cellular processes when compared to humans. Yeast has been used as a model organism to identify various cellular factors that are required for transcriptional silencing. The proteins encoded by the yeast *SIR* genes, which are responsible for maintaining silent chromatin, have been identified as important cellular factors (Shore *et al.*, 1984). It was observed that *SIR2*, *SIR3* and *SIR4* are required for silencing at the mating-type loci (Rine and Herskowitz, 1987) and telomeres (Gottschling *et al.*, 1990), but only *SIR2*, is required for silencing in the rDNA (Bryk *et al.*, 1997, Smith and Boeke, 1997). Silencing of chromatin results in a more closed, inaccessible regional structure. In an active chromatin, specific lysine residues in the amino-terminal tail of histones H3 and H4 are acetylated. The deacetylation of these lysines silences the chromatin which results in the

histones folding into a compact closed nucleosome structure (Braunstein *et al.*, 1996, Hecht *et al.*, 1995, Thompson *et al.*, 1994).

Studies have shown that overexpression of *SIR2* leads to global deacetylation of yeast histones (Braunstein *et al.*, 1993). Thus, *SIR2* was suggested to be a histone deacetylase. Initial attempts to demonstrate histone deacetylase activity were however unsuccessful (Blander and Guarente, 2004). *SIR2*, unlike *SIR3* and *SIR4*, is highly conserved in a wide range of organisms and consequently studies on a bacterial homolog *CobB* were initiated (Brachmann *et al.*, 1995). It was revealed that this gene could substitute for C*obT*, a bacterial gene that encodes an enzyme responsible for transfer of a ribose-phosphate moiety from nicotinic acid mononucleotide (NMN) to dimethyl benzimidazole in the pathway of cobalamin synthesis (Tsang and Escalante-Semerena, 1998). As the cobalamin pathway is not present in yeast or mammals, it was suggested that Sir2 might catalyse a closely related reaction. The adenosine diphosphate (ADP)-riboyltransferase activity of Sir2 proteins from bacteria, yeast and mammals to transfer an ADP-ribose (ADPR) group from nicotinamide adenine dinucleotide (NAD^+) to a protein carrier was reported (Frye, 1999). This activity was shown to be essential for chromatin silencing *in vivo* (Tanny *et al.*, 1999). Additionally, it was shown that the amino-terminal tails of histones H3 and H4 peptides can accept ADPR from NAD^+ when the peptides are acetylated (Imai *et al.*, 2000). However, mass spectrometry was used to confirm that the major function of Sir2 proteins is deacetylation and not ADP-ribosylation. This deacetylation activity of Sir2 proteins is NAD^+-dependent and it was shown to occur only in the presence of NAD^+ and not NADH, NADP and NADPH. These studies led to Sir2 proteins being classified as NAD^+-dependent histone deacetylases (Imai *et al.*, 2000, Smith *et al.*, 2000).

Originally described as a silencing factor (***s***ilent mating type ***i***nformation ***r***egulator **2**) in yeast, Sir2 came on top in a screen for modulators of yeast lifespan (Kaeberlein *et al.*, 1999). Additionally, Sir2 is required for the lifespan of yeast to be extended by calorie restriction (Lin *et al.*, 2000). This led to investigation of Sir2 and its homologues in mammals called sirtuins. The

catalytic mechanism of sirtuins and its role in various diseases are discussed in detail below.

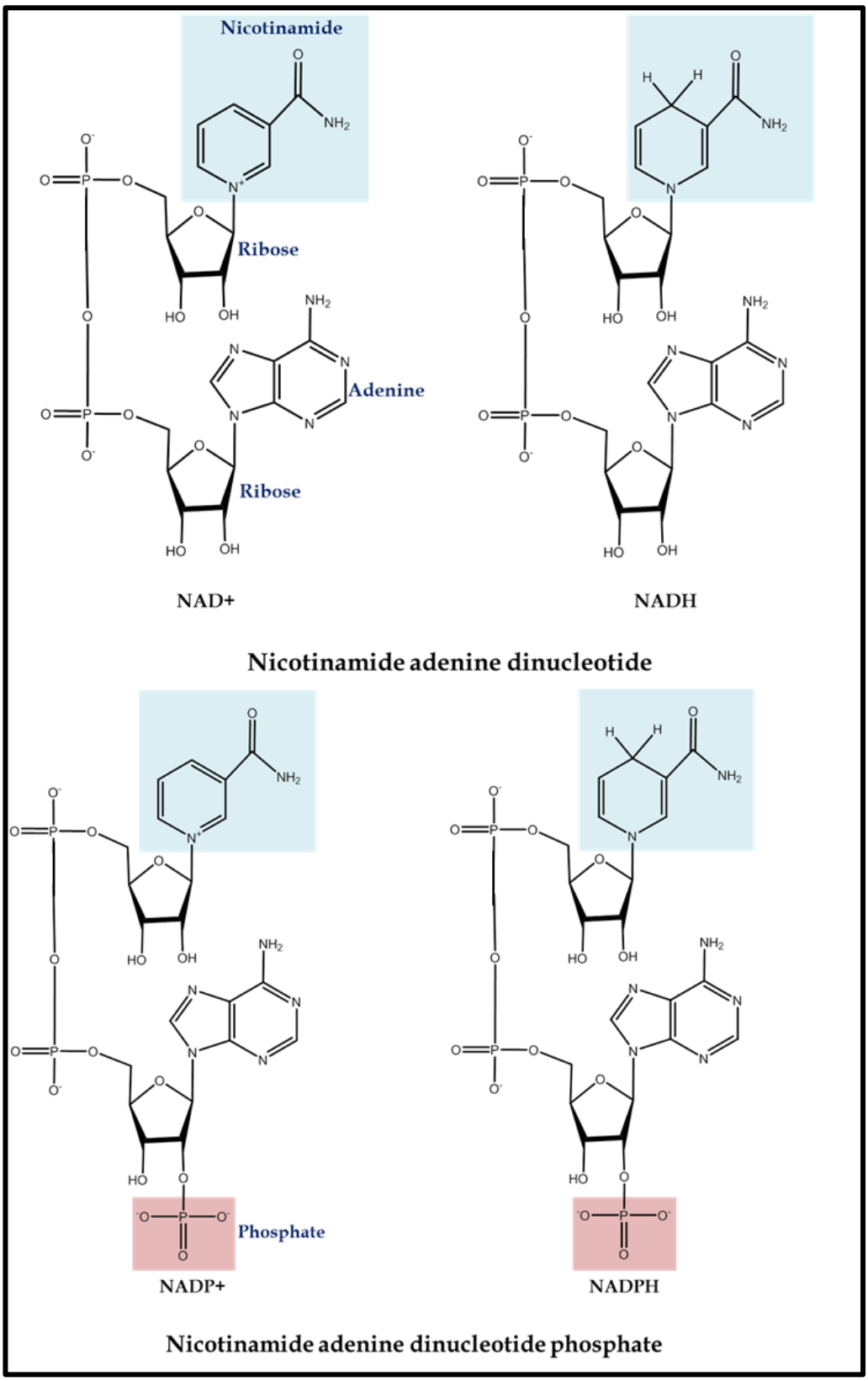


Figure 2. 1: Chemical structures of NAD^+, NADH, $NADP^+$, NADPH

2.1.2. Classification and localisation of Sirtuins

Histone deacetylases (HDACs) are coded for by 18 genes and are grouped into classes I-IV based on their homology to their respective yeast homologues. Classes I, II and IV are termed 'classical' HDACs and comprise 11 members (Witt *et al.*, 2009). They are zinc dependent enzymes that harbour a catalytic domain with a Zn^{2+} ion at its base that can be inhibited by chelating agents such as hydroxamic acids. Sirtuins are class III HDACs which require NAD^{+} as an essential cofactor. Seven mammalian sirtuins have been identified to date. They are defined by their conserved core domain, and some contain additional N- and/or C-terminal sequences [Figure 2.2].

Based on the molecular phylogenetic analysis of 60 sirtuin core domain sequences from prokaryotes and eukaryotes, eukaryotic Sir2-like proteins are grouped into four main branches (Frye, 2000) as shown in Table 2.1. Prokaryotic sirtuins include members of classes II and III. *Sacchromyces cerevisiae* has five class I sirtuins. The seven human sirtuin genes include all four classes: SIRT1, SIRT2 and SIRT3 are class I, SIRT4 is class II, SIRT 5 is class III and SIRT6 and SIRT7 are class IV. A fifth class of sirtuin is present in several gram positive bacteria as well as in the gram negative hyperthermophilic bacterium *Thermatoga maritima*.

Mammalian sirtuins are widely expressed and located in different cellular compartments [Figure 2.3]. Sirtuins in the same compartment, such as mitochondrial SIRT3, SIRT4 and SIRT5 have different sequences and have unique cellular functions and interact with diverse targets (Frye, 2000, Haigis and Guarente, 2006, Michan and Sinclair, 2007). Sirtuins contain nuclear localisation signals (NLSs) as well as nuclear export signals (NESs). Their intracellular localisation is determined by cell/ tissue type and physiological conditions. SIRT1 contains 2 NLS and 2 NES domains while SIRT6 and SIRT7 each contain a single NLS domain (Liszt *et al.*, 2005, Tanno *et al.*, 2007). SIRT3, SIRT4 and SIRT5 contain N-terminal mitochondrial targeting sequences which ensure their localization within the mitochondrial matrix (Onyango *et al.*, 2002, Schwer *et al.*, 2002). This signal sequence is cleaved in the mitochondria which

activates the enzymatic function of the protein. However, under stress, SIRT3 can translocate from the mitochondria to the nucleus (Nakamura *et al.*, 2008, Scher *et al.*, 2007).

Table 2. 1: Classification of sirtuins.

Class	Subclass	Species				Intracellular location
		Bacteria	Yeast	Mouse	Human	
I	A		Sir2p, Hst1	Sirt1	SIRT1	Nucleus, Cytoplasm
	B		Hst2	Sirt2	SIRT2	Cytoplasm
				Sirt3	SIRT3	Nucleus, Cytoplasm, Mitochondria
	C		Hst3, Hst4			
II				Sirt4	SIRT4	Mitochondria
III				Sirt5	SIRT5	Mitochondria
IV	A			Sirt6	SIRT6	Nucleus
	B			Sirt7	SIRT7	Nucleus
U		CobB				

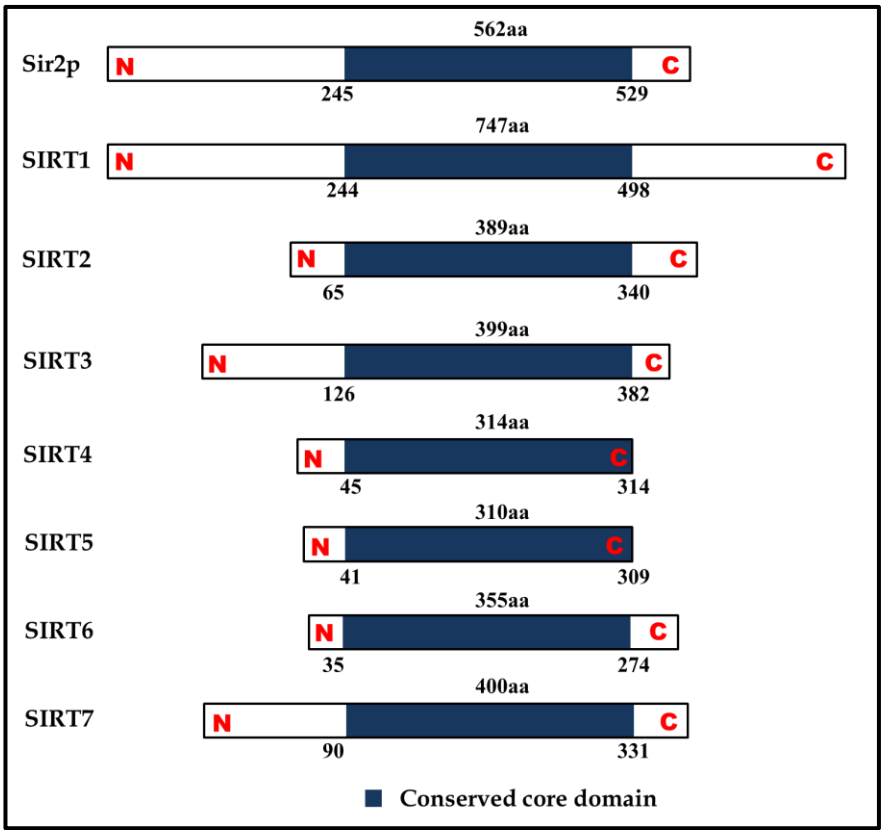


Figure 2. 2: Primary structure of sirtuins.

The seven mammalian sirtuins (SIRT1-7) are aligned with yeast Sir2 (yeast has four other SIR2 paralogs HST1-4). Sizes range from SIRT5 at 33.9 kDa to SIRT 1 at 81.7 kDa. The conserved, catalytic core domain that all sirtuins have in common is in blue. Numbers refer to amino acid residues in the proteins.

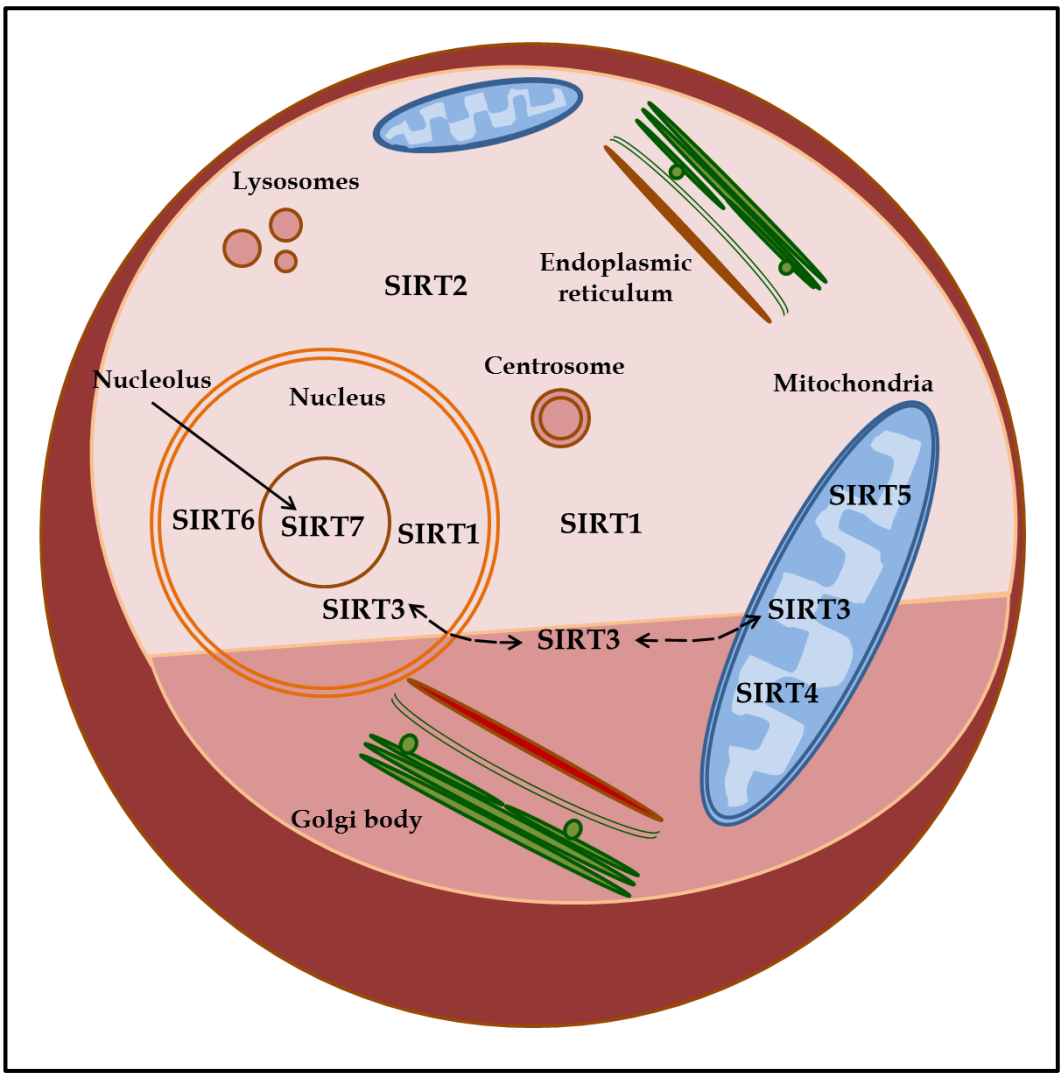


Figure 2. 3: Sirtuin subcellular localization.

SIRT1 is predominantly present in the nucleus and also in the cytosol. SIRT2 is located in the cytosol. SIRT3, SIRT4 and SIRT5 are mitochondrial proteins. However, SIRT3 is also found in the nucleus and cytosol under different cellular events. SIRT6 and SIRT7 are localized in the nucleus and nucleolus, respectively.

2.2. Catalytic mechanism of sirtuins

The Sir-2 dependent deacetylation reaction is different from other classical HDACs due to their NAD^+ dependency. They dynamically regulate transcription, metabolism and cellular responses through deacetylation/deacylation of various substrates. The members of the sirtuin family have a highly conserved common catalytic domain which allows these proteins to recognise a ubiquitously present, metabolically central and abundant compound, NAD^+. The most widely studied reaction catalysed by sirtuins is the NAD^+ dependent protein deacetylation of acetyllysine. For each acetyl lysine that becomes deacetylated, one molecule of NAD^+ is cleaved and effects an acetyl group transfer to ADPR to form 2′-O-acetyl-ADPR (2′-AADPR) (Landry *et al.*, 2000, Tanny and Moazed, 2001). The reaction additionally produces deacetylated lysine residues and nicotinamide as reaction products (Jackson and Denu, 2002). Thus, sirtuins have two coupled enzymatic activities, deacetylation and NAD^+ breakdown [Figure 2.4]. Kinetic and biochemical studies show that the acetyl lysine substrate binds to the enzyme prior to NAD^+. Nicotinamide is cleaved from NAD^+ and is the first product released, followed by deacetylated lysine and 2′-AADPR. In aqueous solution, spontaneous non-enzymatic equilibration of AADPR isomers, 2′- and 3′-AADPR occurs. Several AADPR metabolising enzymes have been reported, including nucleoside diphosphate linked to moiety x (NUDIX) hydrolases, ADP-ribosyl hydrolase-3 (ARH3) and Macro domain proteins (Tong and Denu, 2010). NUDIX hydrolases cleaves the pyrophosphate bond of AADPR forming AMP and 2- and 3- O-acetylribose-5-phosphate. Human Macro domain proteins hydrolyse AADPR to form free acetate and ADPR. In lower organisms, sirtuins and Marco domain proteins are present as fusion proteins or are genetically connected via the same operon, suggesting that another pathway may involve AADPR, sirtuins and Macro domain proteins. AADPR is a less explored metabolite that exhibits signalling functions and is suggested to act as a substrate for downstream enzymatic processes. It has been linked with gene silencing, ion-channel activation and decreased reactive oxygen species (Tong and Denu, 2010).

The first step catalysed by sirtuins leading towards acetyl-lysine deacetylation is ADP-ribosylation of acetyllysine which results in the formation of an ADP-ribosyl-peptidylimidate. Various mechanisms have been proposed for the formation of this species. Three main mechanisms have been shown in figure 2.5 which include S_N1, concerted S_N2 or dissociative asynchronous mechanisms (Sauve, 2010). Data in support of each of these mechanisms have been produced, although there is a lack of firm consensus on which of the mechanisms best explains sirtuin chemistry (Sauve, 2010).

The characterised general nucleophiles for NAD^+ include acetyl lysine substrates or thioacetyl lysine inhibitors (Sauve, 2010). The formation of an akylamidate intermediate is supported by biochemical studies (Feldman *et al.*, 2012). Upon formation of the C1′-O-akylamidate intermediate, the 2′- hydroxyl group of NAD^+ ribose is activated by a conserved active-site histidine. The activated hydroxyl attacks the O-akylamidate carbon to form a 1′2′-cyclic intermediate. A base activated water molecule attacks the cyclic intermediate to form deacetylated lysine and AADPR [Figure 2.6].

In addition to the deacetylase activity, certain sirtuins also catalyse other reactions such as protein ADP-ribosyl transfer, NAD^+ hydrolysis and many if not all catalyse acetyl lysine-dependent nicotinamide base exchange into NAD+ (Sauve, 2010). Understanding the mechanistic details of sirtuin-catalysed reactions is an important step towards a complete understanding of sirtuin functions. This in turn would help the development of novel chemical tools to probe their biology.

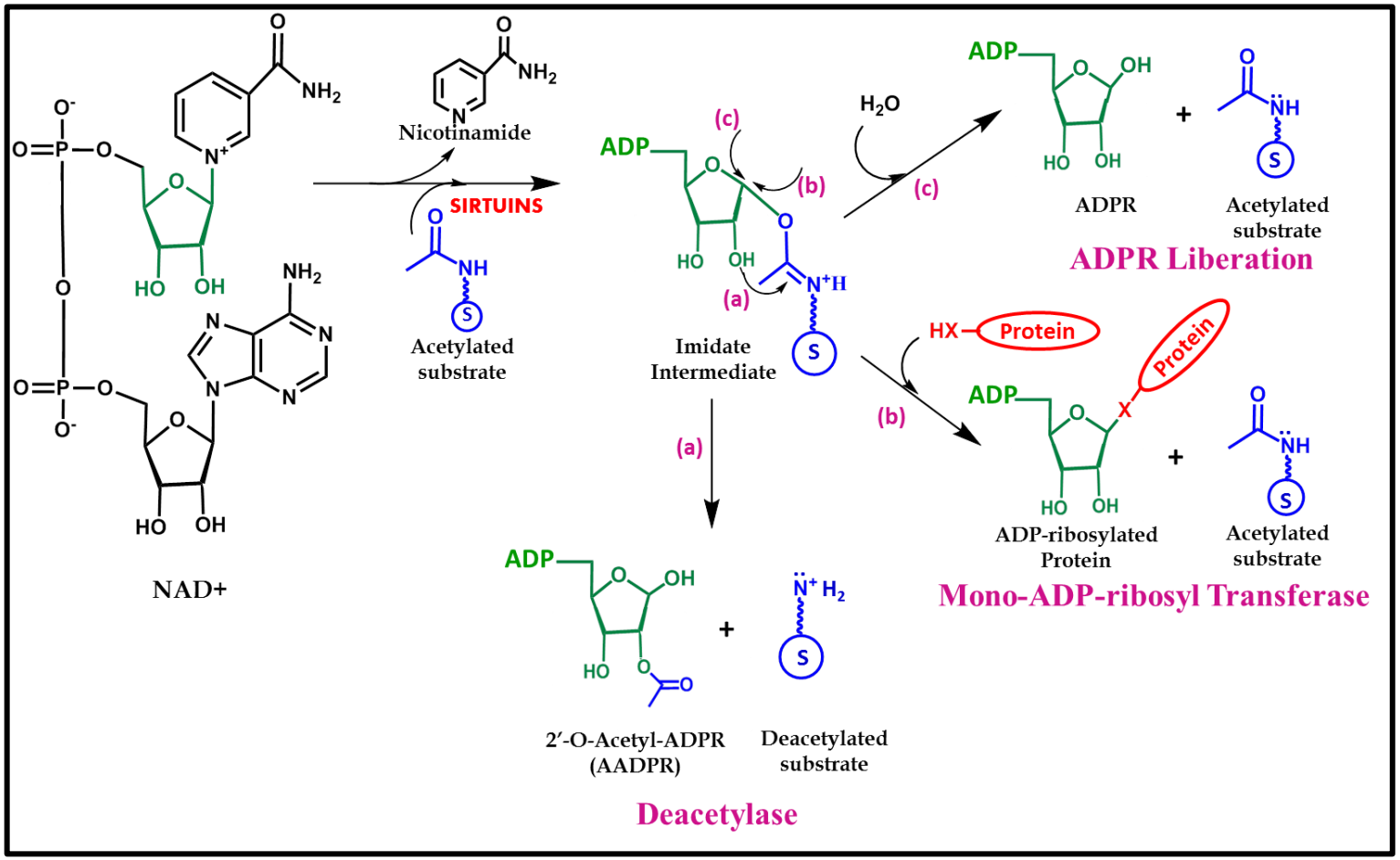


Figure 2. 4: Sirtuin catalysed reactions with a common intermediate.

The key intermediate is an imidate which can be processed into different products depending on the attacking nucleophiles. **(a)Deacetylation:** The major reaction in the presence of acetylated substrates. **(b) ADP-ribosylation** and in some cases self ribosylation of sirtuins. **(c) Liberation of ADPR**. Adapted from (Du *et al.*, 2009, Sauve, 2010)

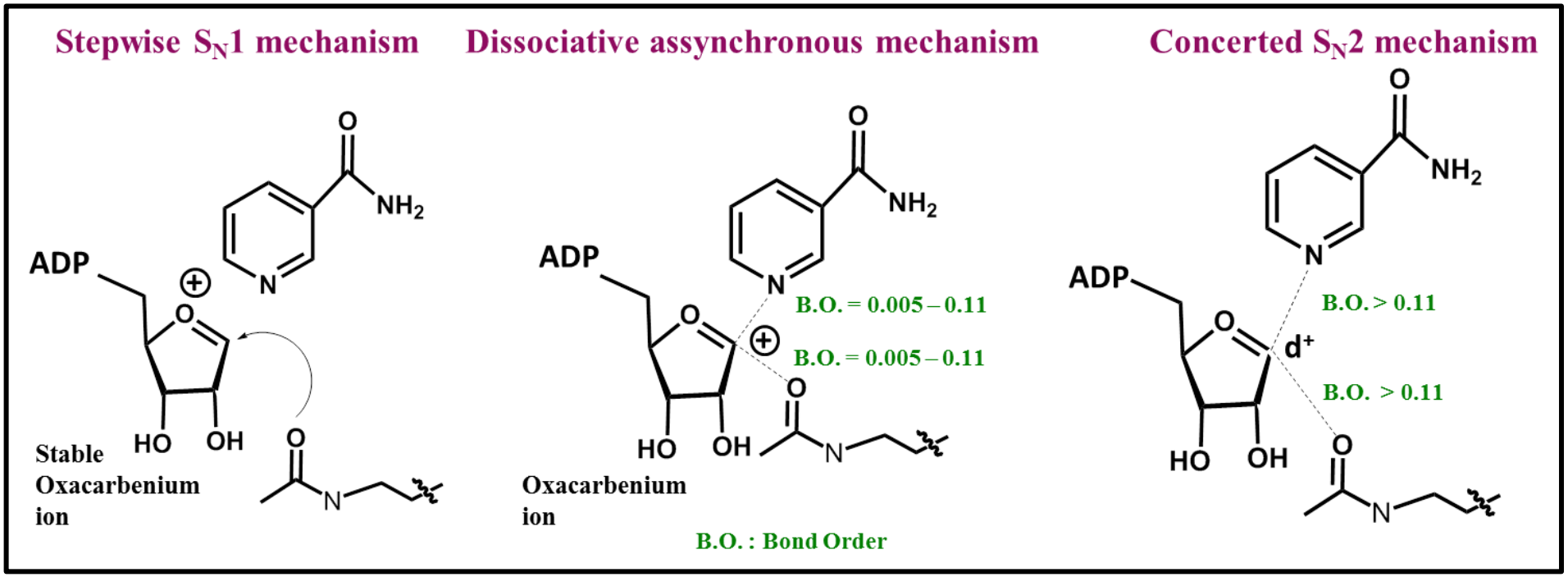


Figure 2. 5: Proposed transition states for ADP-ribosylation of acetyl-lysine.

(a) S_N1 mechanism: Complete dissociation of nitocinamide takes place and an enzyme-stabilising oxacarbenium ion is formed, followed by collapse of the oxacarbenium ion by reaction with acetyllysine. **(b)Dissociative asynchronous mechanism:** Oxacarbenium ion is formed at the transition state but with extensive bond cleavage to the leaving group and weak bond formation to the nucleophile. This reaction is not stepwise as it lacks a discrete intermediate between the reactants and the imidate. As the leaving group is largely cleaved prior to bond formation with the nucleophile, this reaction is asynchronous. **(c) S_N2 mechanism:** A penta-coordinate transition state with significant bond orders to both the nucleophile and the leaving group is formed. Hence, this mechanism is characterised by associative effects of the nucleophile at the transition state, consistent with an S_N2 mechanism. Adapted from (Sauve, 2010)

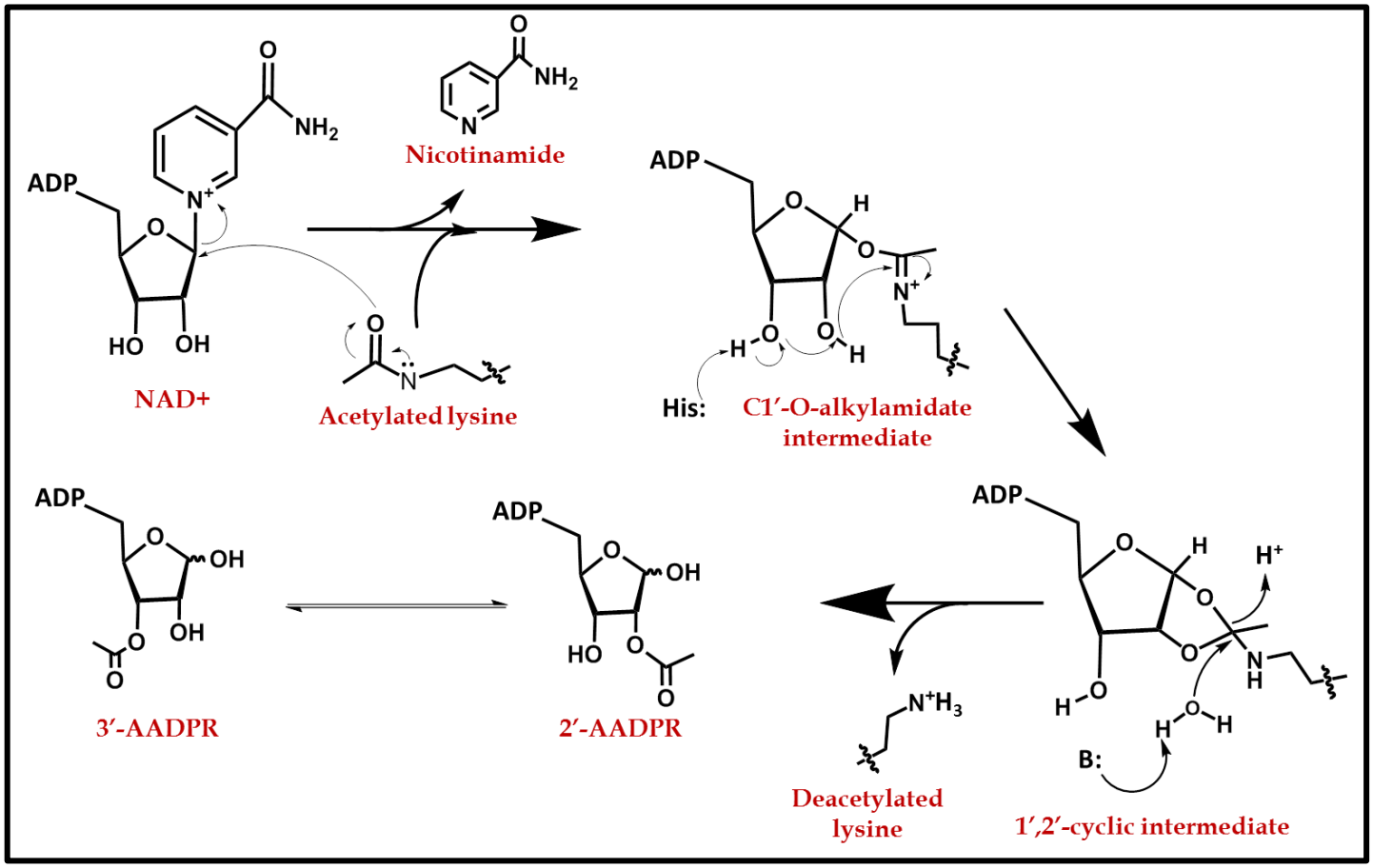


Figure 2. 6: Mechanism of sirtuin catalysed deacetylation

Nucleophilic addition of the acetyl oxygen to the 1′-carbon of the nicotinamide ribose affords alkylamidate intermediate. The 2′-OH is activated by a conserved histidine residue that attacks the alkylamidate carbon to afford a cyclic intermediate. A base activated water molecule degrades the intermediate forming deacetylated lysine and AADPR. (Feldman *et al.*, 2012, Sauve, 2010)

2.3. Sirtuin activity and their targets

Sirtuins are involved in regulation of a wide range of processes in normal cellular physiology through NAD^+-dependent deacetylation /deacylation of various substrates. Lysine acetylation is a global post-translational modification that regulates diverse cellular processes and has been extensively studied (Feldman *et al.*, 2012). A number of sirtuin targets have been identified and the number is continually increasing [Table 2.2, Figure 2.7].

Table 2. 2: Substrates of sirtuins

Modified from (Blander and Guarente, 2004, Feldman *et al.*, 2012, Houtkooper *et al.*, 2012, McGuinness *et al.*, 2011)

Sirtuin	Class	Activity	Targets
SIRT1	I	Deacetylation	Histones (H1, H2, H3 and H4) proliferator-activated receptor γ coactivator-1α (PGC1α), Forkhead box (FOXO), p53, Notch, Nuclear factor- κB (NF-κB), Hypoxia inducible factor 1α (HIF1α), Liver X receptor (LXR), Farnesoid X receptor (FXR), Sterol-response element-binding protein 1c (SREBP1c), p300, Ku70, Peroxisome proliferator-activate receptor γ (PPARγ), Trans-activator of transcription (Tat), Estrogen receptor (ERα), Androgen receptor (AR), Mothers against decapentaplegic homolog 7 (SMAD7), p73, Sex determining region Y- box 9 (Sox9), Hairy /enhancer of split-1 (HES1), Hairy/enhancer of split related with YRPW motif protein 2 (HEY2), Nuclear receptor co-repressor 2/ Thyroid hormone receptor (NcoR/SMRT), p65, Phosphoglycerate mutase 1 (PGAM-1), Acetyl-CoA synthase 1 (AceCS1), among others
SIRT2	I	Deacetylation	α-tubulin, FOXO, Phosphoenolpyruvate carboxykinase (PEPCK), Protease activated receptor (PAR3), H4

SIRT3	I	Deacetylation	Long-chain acyl CoA dehydrogenase (LCAD), 3-hydroxy-3-methylglutaryl CoA synthase 2 (HMGCS2), Glutamate dehydrogenase (GDH), Oxidative phosphorylation (OXPHOS) complexes, Superoxide dismutase 2 (SOD2), Isocitrate dehydrogenase 2 (IDH2), Acetyl-CoA synthase 2 (AceCS2), Ornithine transcarbamylase (OTC), Succinate dehydragenase (SDH)
SIRT4	II	ADP-ribosylation	GDH, Insulin degrading enzyme (IDE), Adenine nucleotide translocator (ANT)
SIRT5	III	Deacetylation, demalonylation, desuccinylation	Carbamoyl phosphate synthetase 1 (CPS1), Superoxide dismutase 2
SIRT6	IV	Deacetylation, ADP-ribosylation	H3, NF-κB, HIF1α, helicase, DNA polymerase β, Choline transporter-like protein (CtlP), Poly[ADP-ribose] polymerase 1 (PARP1), PPARγ
SIRT7	IV	Deacetylation	RNA polymerase type I, Adenovirus early region 1A (E1A), SMAD6

SIRT1 is an NAD^+-dependent deacetylase localised in the nucleus (Imai *et al.*, 2000). SIRT1 deacetylates a number of histone and non-histone proteins including histone H3 and H4, p53, NF-κB, PGC-1α, PPARγ, MyoD (a protein involved in regulating muscle differentiation) and HIF1α, among others (Blander and Guarente, 2004, Feldman *et al.*, 2012, Houtkooper *et al.*, 2012, McGuinness *et al.*, 2011). These targets imply SIRT1 plays an important role in glucose metabolism, fatty acid and cholesterol metabolism, differentiation, insulin secretion and neuroprotection.

SIRT1 plays a vital role in the regulation of homeostasis, cell metabolism, cell signalling and stress responses. It can affect the core histones (H1, H2, H3 and H4) but it preferentially deactylates H3 (K9, K14 and K56), H4 (K16) and H1 (K26) (McGuinness *et al.*, 2011).The deacetylation of lysine residues at H3K9, H4K16 and H1K26 have been associated with gene silencing and chromatin remodelling. Deacetylation of histones facilitates its methylation which has been linked to global transcriptional repression which is enhanced by SIRT1

(Bartova *et al.*, 2008). SIRT1 binds to histone-lysine N-methyltransferase Suv39H1 (suppressor of variegation 3-9 homologue 1) and enables binding of this protein to chromatin. SIRT1 is then activated by deacetylation of Suv39H1 (McGuinness *et al.*, 2011, Vaquero *et al.*, 2007). DBC1 (deleted in breast cancer 1) inhibits SIRT1 and Suv39H1 and disrupts the interaction between these two molecules which increases the methylation of H3K9. This, in turn, regulates the chromatin silencing mediated by SIRT1-Suv39H1 complex (Kim *et al.*, 2008). SIRT1 is also involved in the epigenetic regulation of gene expression. This is due to its interaction with DNA methyltranferases, DNMT3 and DMNT1, which results in aberrant expression of methylated genes (Espada *et al.*, 2007, O'Hagan *et al.*, 2008). The ability of SIRT1 to deacetylate and interact with a broad range of transcription factors [Table 2.2] and its ability to remodel chromatin, suggests that SIRT1 plays a vital role in the regulation of organism homeostasis, stress responses, endocrine signalling and cell metabolism.

SIRT2 deacetylation targets include acetylated lysine residues H4K16 and H3K9 when the nuclear envelope disassembles during the mitosis (Vaquero *et al.*, 2006). Thus, SIRT2 acts as a regulator of the cell cycle and is involved in the promotion of chromatin condensation. SIRT2 is known to co-localise with microtubules, deacetylate K40 of α-tubulin and SIRT2 RNA*i* results in hyperacetylation of tubulin (North *et al.*, 2003).

SIRT3-5 localise in the mitochondria and play important roles as metabolic sensors that respond to changes in the energy status of the cell. They are involved in the modulation of activity of key metabolic enzymes via protein deacetylation (Martinez-Pastor and Mostoslavsky, 2012, Newman *et al.*, 2012). SIRT3 is also involved in ATP production and fatty acid oxidation. The roles of SIRT4 in insulin secretion and SIRT5 in the urea cycle have also been established (McGuinness *et al.*, 2011).

SIRT6 is involved in the deacetylation of H3K9 at the telomeres suggesting its involvement in the modulation of cellular senescence and ageing induced chromosomal abnormalities (Michishita *et al.*, 2009).

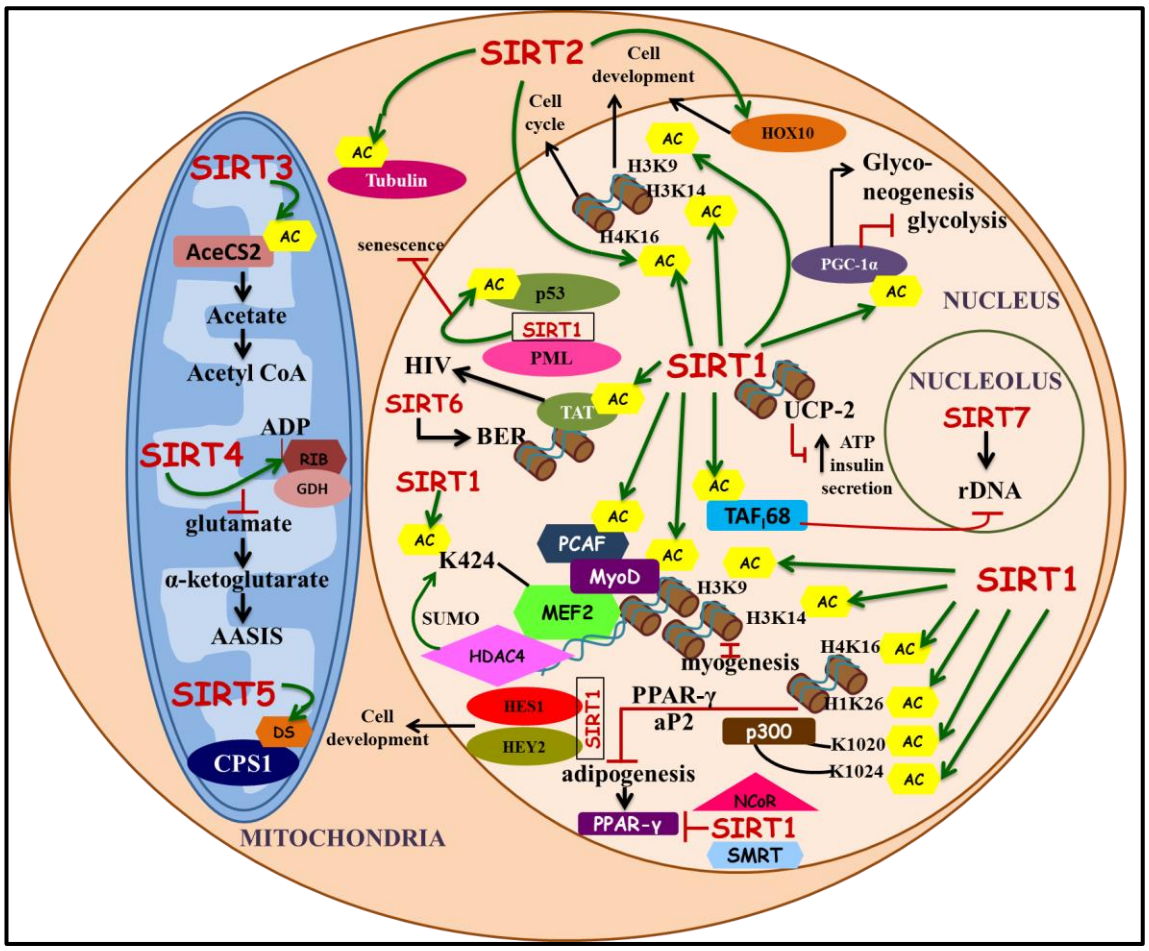


Figure 2. 7: Sirtuins in mammalian cells.

Sirtuins regulate a wide range of processes in the human cells. In the nucleus, SIRT1 modulates chromatin structure by deacetylating specific lysine residues in histones and alters gene expression by targeting lysine residues of p300. It also promotes HIV-1 replication by deacetylating Tat. SIRT1 protects cells from senescence by complexing with PML and mediating p53 deacetylation. In complexing with PCAF and by deacetylating K424 of MEF2, SIRT1 supresses muscle differentiation. SIRT1 inhibits adipogenesis by interacting PPAR-γ and aP2 promoters, as well as with PPAR-γ co-factors NcoR and SMRT. SIRT1 regulates cell development by interacting with HES1/HEY2. SIRT2 also helps in development along with homebox transcription factor HOX10. Deacetylation of PCG-1α modulates glucose metabolism. SIRT6 localises to heterochromatin to regulate base excision repair. SIRT2 translocates from the cytoplasm to the nucleus during mitosis and deacetylates histone H4 K16. In the nucleolus, SIRT7 activates RNS polymerase I transcription while SIRT1 inhibits it by deacetylating TAFI68. Mitochondrial SIRT3 deacetylates and activates AceCS2 while SIRT 4 mono-ADP-ribosylates GDH which in turn inhibits amino acid stimulates insulin secretion (AASIS). SIRT5 desuccinylates CPS1 to activate the urea cycle. Modified from (Michan and Sinclair, 2007)

Sirtuins are associated with the regulation of RNA polymerase II transcribed genes and the transcription of ribosomal RNA (McGuinness *et al.*, 2011). SIRT1 deacetylates TATA box-binding protein associated factor (TAF_168) leading to the inhibition of polymerase I (Muth *et al.*, 2001). In contradiction to this role of SIRT1 on rRNA transcription, SIRT7 directly binds to polymerase I and induces enzymatic activity (Ford *et al.*, 2006). Decrease in SIRT7 levels stops cell proliferation and triggers apoptosis, thus SIRT7 plays an important role in cell viability in mammals.

In addition to the deacetylation role of sirtuins, SIRT1-3 can catalyse depropionylation and debutyrylation but with varying efficiencies compared to deacetylation (Garrity *et al.*, 2007, Smith and Denu, 2007). The ability of SIRT5 to catalyse desuccinylation and demalonylation rather than deacetylation has also been established (Du *et al.*, 2011). Deletion of SIRT5 has shown to increase the level of succinylation on carbamoyl phosphate synthase 1 (CPS1), a known target of SIRT5 (Du *et al.*, 2011).

All sirtuin targets identified to date reveal the involvement of sirtuins in cellular stress responses which include cell death responses, senescence, stress-related transcription regulation, cell-cycle regulation, cell metabolism, genomic formation and stability, maintenance of telomeric function and its control, among others [Table 2.2]. This understanding of sirtuin targets has unveiled their association with telomeric stability and cellular lifespan, and in turn with longevity. Understanding sirtuins and their functions might reveal the causes of ageing related diseases and help in the design of new tools for their treatment.

2.4. Structural basis for sirtuin function:

The three dimensional structures of several sirtuin homologues in their native form and in complex with other ligands have been determined by X-ray crystallography to aid in understanding their catalytic mechanism, substrate specificity and inhibitory mechanisms. These include sirtuins from archaeon *Archaeoglobus fulgidus*, bacteria *Thermotoga maritima* and *Escherichia coli*, yeast *Sacchaaromyces cerevisiae* and several human sirtuins (Bernstein *et al.*, 1977). This wealth of structural information together with complementary biochemical studies has provided valuable insights into sirtuin activities. A brief summary of the salient features of the high resolution structures of human sirtuins [Table2.3] is given here.

Table 2. 3: Summary of human sirtuin structures

Protein	PDB ID	Details	Reference
SIRT1	4KXQ	SIRT1 (residues 234-510 and 641 – 665) – ADPR (closed form)	(Davenport *et al.*, 2014)
	4IG9	SIRT1 (residues 234-510 and 641 – 665) native (open form)	
	4I5I	SIRT1 (residues 241-516) - NAD^+ - indole analogue of EX527 (inhibitor)	(Zhao *et al.*, 2013)
SIRT2	3ZGV	SIRT2 –ADPR	(Moniot *et al.*, 2013)
	3ZGO	Apo-form SIRT2	
	1J8F	Native SIRT2	(Finnin *et al.*, 2001)
	4L30	SIRT2 - macrocyclic peptide S2iL5	(Yamagata *et al.*, 2014)
SIRT3	4BVE	SIRT3 – Thioalkylimidate from thio-acetllysine ACS2 peptide	(Gertz *et al.*, 2013)
	4BVB	SIRT3 - ADPR – EX527 (inhibitor)	
	4BVG	SIRT3 with native alkylimidate from acetyllysine ACS2 peptide – EX-527	
	4BVF	SIRT3 - native thioalkylimidate from thio-acetyllysine ACS2 peptide – EX527	
	4BVH	SIRT3 – 2′-AADPR – EX527 (inhibitor)	
	4BV3	SIRT3 - NAD^+ - EX527 (inhibitor)	
	4C78	SIRT3 – Bromoresveratrol- ACS2 peptide	(Nguyen *et al.*, 2013a)
	4C7B	SIRT3 – Bromoresveratrol – Fluor-de-lys	
	4BN4	SIRT3 – ADPR	(Nguyen *et al.*, 2013b)
	4BN5	SIRT3 – SRT1720 (inhibitor)	

	4JSR	SIRT3 – ELT inhibitor 11c	(Disch *et al.*, 2013)
	4JT8	SIRT3 – ELT inhibitor 28	
	4JT9	SIRT3 – ELT inhibitor 3	
	4FZ3	SIRT3 – acetyl p53 peptide coupled with 4-amino-7-methylcoumarin	(Wu *et al.*, 2013)
	3GLS	Native SIRT3	(Jin *et al.*, 2009)
	3GLT	SIRT3 – ADPR – ACS2 peptide with thioacetyllysine	
	3GLU	SIRT3 – ACS2 peptide	
	3GLR	SIRT3 – ACS2 peptide with thioacetyllysine	
	4HD8	SIRT3- Fluor-de-lys peptide and peceatanol	(Gertz *et al.*, 2012)
SIRT5	2NYR	SIRT5 – suramin	(Schuetz *et al.*, 2007)
	3RIG	SIRT5 – Thioacetyl H3K9	(Du *et al.*, 2011)
	3RIY	SIRT5 - Succinyl H3K9 - NAD	
	4F4U	SIRT5 - Succinyl H3K9	(Zhou *et al.*, 2012)
	4F56	SIRT5 – Bicyclic intermediate	
	4HDA	SIRT5-Fluor-de-lys - Resveratrol	(Gertz *et al.*, 2012)
	4G1C	SIRT5 – Sicciny IDH2 and Carba-NAD	(Szczepankiewicz *et al.*, 2012)
SIRT6	3PKI	SIRT6 ADPR	(Pan *et al.*, 2011)
	3PKJ	SIRT6 2'-N-acetyl ADPR	

Primary sequence alignment [Figure 2.8] and ConSurf analysis [Figure 2.9] of sirtuin proteins indicate that they contain a conserved catalytic core of approximately 270 residues. Additionally, the sirtuin proteins have N- and C-terminal flanking regions that are variable in length and sequence and have various roles in cellular localisation, oilgomerisation and autoregulation mechanisms (Brachmann *et al.*, 1995, Frye, 1999, Tennen *et al.*, 2010). Unlike other protein deacetylase families that activate a water molecule for hydrolysis reactions by using zinc as a cofactor, sirtuin mechanisms utilise NAD^+ as a co-substrate.

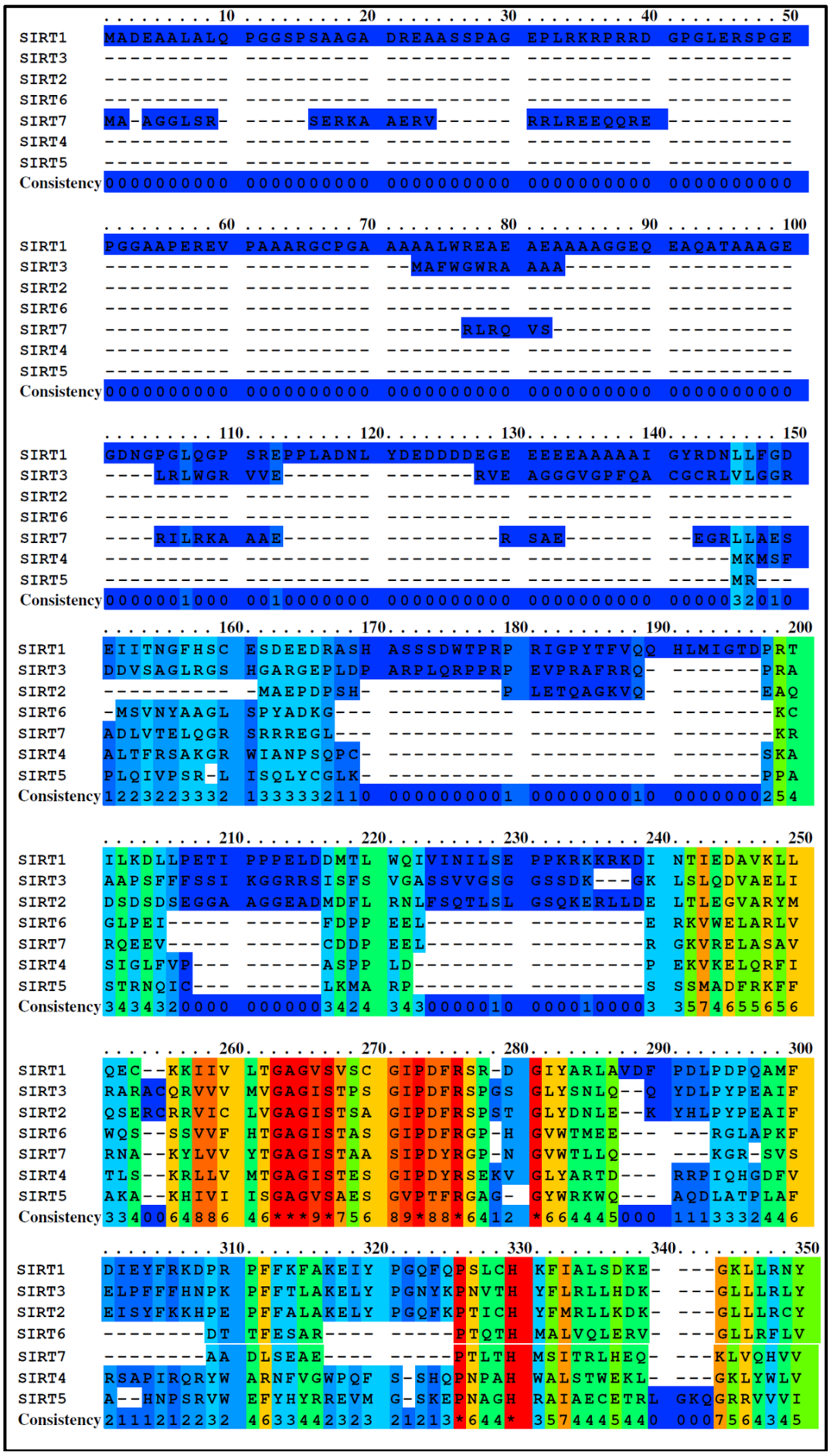
The catalytic core region of sirtuins adopts an elongated oval shaped fold composed of two globular subdomains linked by several loops which contribute to the active site cleft between the subdomains [Figure 2.10A]. The

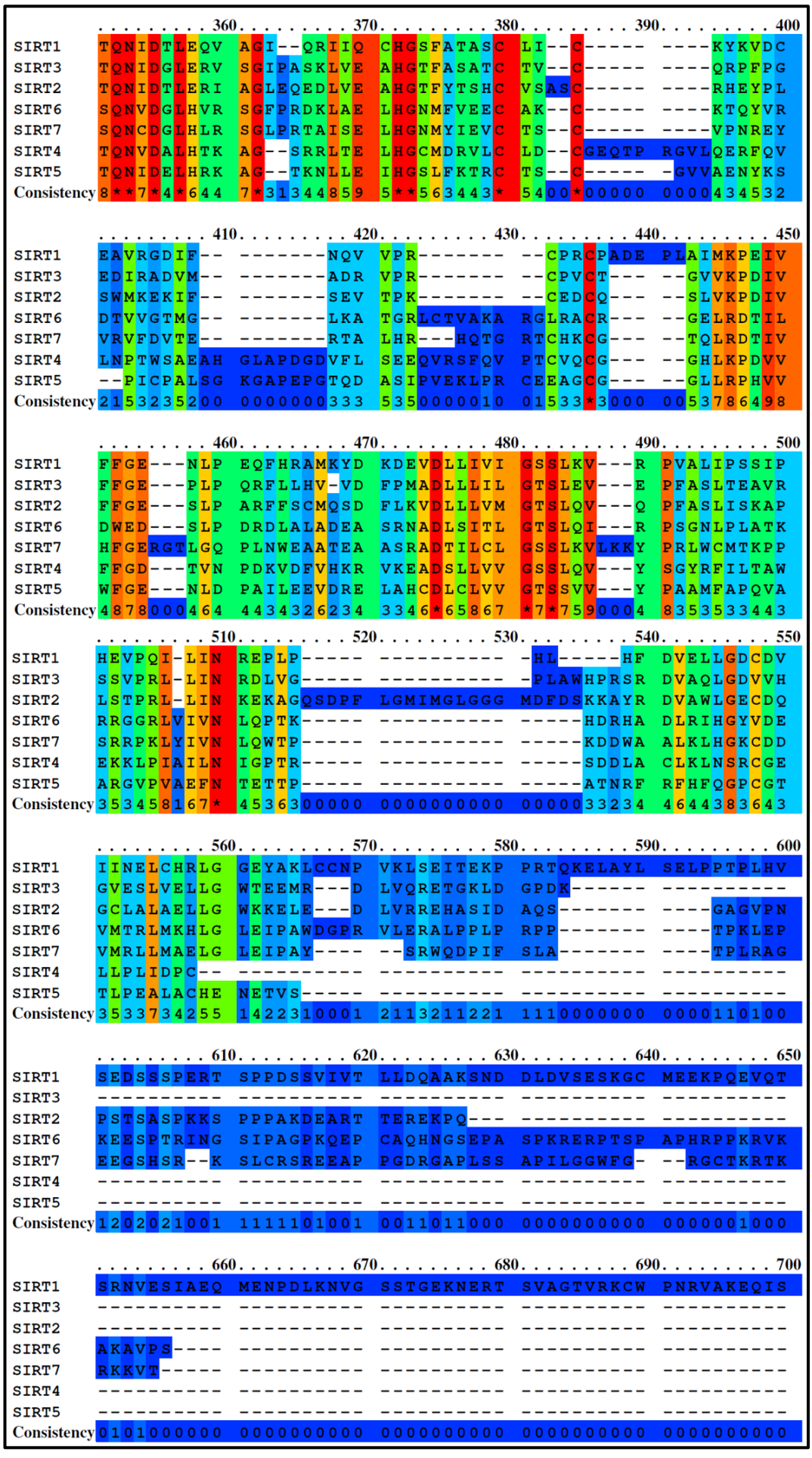
two globular subdomains differ in size. The larger domain is highly conserved and consists of a Rossmann-fold domain, which is characteristic of NAD^+/NADH binding proteins. The Zn-binding motif, bearing the sequence Cys-X_{2-4}-Cys-X_{15-40}-Cys-X_{2-4}-Cys, and an α-helical region, show the highest diversity among members of the sirtuin family. The loops connecting the two domains form a pronounced, extended cleft between the two domains where the NAD^+ and the acetyl-lysine containing peptide substrate enter from opposite sides and bind to the enzyme [Figure 2.10]. The NAD^+ binding cleft is divided into three regions: site A for adenine binding, site B for nicotinamide-ribose binding and site C for nicotinamide moiety binding [Figure 2.10C].

The relative orientations of the small and large domains vary in known sirtuin structures and may be attributed to the presence or absence of substrates at the substrate binding cleft. The co-factor binding loop shows a high degree of flexibility and its conformation was shown to change in close relation to the catalytic events during the formation of a covalent 1′-O-alkylamidate intermediate between the two substrates under nicotinamide release (Sanders *et al.*, 2010). This is followed by the hydrolysis of the intermediate to yield deacetylated polypeptide and 2′-AADPR. The amino acids involved in the catalysis and the reactive groups of both bound substrate molecules are buried within a protein tunnel in the cleft between the two domains which is the region with a high degree of conservation. The cofactor binding loop becomes ordered upon substrate binding and changes to a more closed conformation upon acetyl transfer (Chang *et al.*, 2002, Hoff *et al.*, 2006). This loop-relocation suggests that the cofactor binding loop may be involved in ejecting the first reaction product, nicotinamide. The rearrangements of the co-factor binding loop and the acetyllysine binding tunnel highlight the dynamic structure of the sirtuins. The differences in activity and affinity to substrates among the members of the sirtuin family is attributed to the differences in shape and electrostatics of the peptide binding groove (Sanders *et al.*, 2010).

The structural and biochemical studies to date have provided insights into the activity of the sirtuins. However, to answer specific question such as the

difference in the deacetylation and ribosylation activity among mammalian sirtuins, the biological role of regions N- and C-terminal to the catalytic core and inhibition mechanisms of synthetic sirtuin inhibitors will require the structure determination of sirtuin complexes containing reaction intermediates at various stages of the reaction. The dynamic nature, wide range of substrate recognition and adaptability of the members of the sirtuin family highlight the importance of obtaining a reliable picture for understanding their activity and for rational drug development.


10 20 30 40 50
SIRT1 MADEAALALQ PGGSPSAAGA DREAASSPAG EPLRKRPRRD GPGLERSPGE
SIRT3 ---------- ---------- ---------- ---------- ----------
SIRT2 ---------- ---------- ---------- ---------- ----------
SIRT6 ---------- ---------- ---------- ---------- ----------
SIRT7 MA-AGGLSR- -----SERKA AERV------ RRLREEQQRE ----------
SIRT4 ---------- ---------- ---------- ---------- ----------
SIRT5 ---------- ---------- ---------- ---------- ----------
Consistency 0000000000 0000000000 0000000000 0000000000 0000000000
60 70 80 90 100
SIRT1 PGGAAPEREV PAAARGCPGA AAAALWREAE AEAAAAGGEQ EAQATAAAGE
SIRT3 ---------- ---------- --MAFWGWRA AAA------- ----------
SIRT2 ---------- ---------- ---------- ---------- ----------
SIRT6 ---------- ---------- ---------- ---------- ----------
SIRT7 ---------- ---------- ------RLRQ VS-------- ----------
SIRT4 ---------- ---------- ---------- ---------- ----------
SIRT5 ---------- ---------- ---------- ---------- ----------
Consistency 0000000000 0000000000 0000000000 0000000000 0000000000
110 120 130 140 150
SIRT1 GDNGPGLQGP SREPPLADNL YDEDDDDEGE EEEEAAAAAI GYRDNLLFGD
SIRT3 ----LRLWGR VVE------- -------RVE AGGGVGPFQA CGCRLVLGGR
SIRT2 ---------- ---------- ---------- ---------- ----------
SIRT6 ---------- ---------- ---------- ---------- ----------
SIRT7 ----RILRKA AAE------- ---------R SAE------- --EGRLLAES
SIRT4 ---------- ---------- ---------- ---------- -----MKMSF
SIRT5 ---------- ---------- ---------- ---------- -----MR---
Consistency 0000001000 0010000000 0000000000 0000000000 0000032010
160 170 180 190 200
SIRT1 EIITNGFHSC ESDEEDRASH ASSSDWTPRP RIGPYTFVQQ HLMIGTDPRT
SIRT3 DDVSAGLRGS HGARGEPLDP ARPLQRPPRP EVPRAFRRQ- -------PRA
SIRT2 ---------- -MAEPDPSH- ---------P LETQAGKVQ- -------EAQ
SIRT6 -MSVNYAAGL SPYADKG--- ---------- ---------- --------KC
SIRT7 ADLVTELQGR SRRREGL--- ---------- ---------- --------KR
SIRT4 ALTFRSAKGR WIANPSQPC- ---------- ---------- -------SKA
SIRT5 PLQIVPSR-L ISQLYCGLK- ---------- ---------- -------PPA
Consistency 1223223332 1333332110 0000000001 0000000010 0000000254
210 220 230 240 250
SIRT1 ILKDLLPETI PPPELDDMTL WQIVINILSE PPKRKKRKDI NTIEDAVKLL
SIRT3 AAPSFFFSSI KGGRRSISFS VGASSVVGSG GSSDK---GK LSLQDVAELI
SIRT2 DSDSDSEGGA AGGEADMDFL RNLFSQTLSL GSQKERLLDE LTLEGVARYM
SIRT6 GLPEI----- ------FDPP EEL-------- ----------E RKVWELARLV
SIRT7 RQEEV----- ------CDDP EEL-------- ----------R GKVRELASAV
SIRT4 SIGLFVP--- ------ASPP LD-------- ----------P EKVKELQRFI
SIRT5 STRNQIC--- ------LKMA RP-------- ----------S SSMADFRKFF
Consistency 3434320000 000000 3424 3430000010 0000100003 3574655656
260 270 280 290 300
SIRT1 QEC--KKIIV LTGAGVSVSC GIPDFRSR-D GIYARLAVDF PDLPDPQAMF
SIRT3 RARACQRVVV MVGAGISTPS GIPDFRSPGS GLYSNLQ--Q YDLPYPEAIF
SIRT2 QSERCRRVIC LVGAGISTSA GIPDFRSPST GLYDNLE--K YHLPYPEAIF
SIRT6 WQS--SSVVF HTGAGISTAS GIPDFRGP-H GVWTMEE--- ---RGLAPKF
SIRT7 RNA--KYLVV YTGAGISTAA SIPDYRGP-N GVWTLLQ--- ---KGR-SVS
SIRT4 TLS--KRLLV MTGAGISTES GIPDYRSEKV GLYARTD--- RRPIQHGDFV
SIRT5 AKA--KHIVI ISGAGVSAES GVPTFRGAG- GYWRKWQ--- AQDLATPLAF
Consistency 3340064886 46***9*756 89*88*6412 *6644450001 1133332446
310 320 330 340 350
SIRT1 DIEYFRKDPR PFFKFAKEIY PGQFQPSLCH KFIALSDKE- ---GKLLRNY
SIRT3 ELPFFFHNPK PFFTLAKELY PGNYKPNVTH YFLRLLHDK- ---GLLLRLY
SIRT2 EISYFKKHPE PFFALAKELY PGQFKPTICH YFMRLLKDK- ---GLLLRCY
SIRT6 --------DT TFESAR---- -----PTQTH MALVQLERV- ---GLLRFLV
SIRT7 --------AA DLSEAE---- -----PTLTH MSITRLHEQ- ---KLVQHVV
SIRT4 RSAPIRQRYW ARNFVGWPQF S-SHQPNPAH WALSTWEKL- ---GKLYWLV
SIRT5 A--HNPSRVW EFYHYRREVM G-SKEPNAGH RAIAECETRL GKQGRRVVVI
Consistency 2111212232 4633442323 21213*644* 3574445440 0007564345


. . . . . 360 . . . . . 370 . . . . . 380 . . . . . 390 400
SIRT1 TQNIDTLEQV AGI--QRIIQ CHGSFATASC LI--C----- ----KYKVDC
SIRT3 TQNIDGLERV SGIPASKLVE AHGTFASATC TV--C----- ----QRPFPG
SIRT2 TQNIDTLERI AGLEQEDLVE AHGTFYTSHC VSASC----- ----RHEYPL
SIRT6 SQNVDGLHVR SGFPRDKLAE LHGNMFVEEC AK--C----- ----KTQYVR
SIRT7 SQNCDGLHLR SGLPRTAISE LHGNMYIEVC TS--C----- ----VPNREY
SIRT4 TQNVDALHTK AG--SRRLTE LHGCMDRVLC LD--CGEQTP RGVLQERFQV
SIRT5 TQNIDELHRK AG--TKNLLE IHGSLFKTRC TS--C----- -GVVAENYKS
Consistency 8**7*4*644 7*313444859 5**563443* 5400*00000 0000434532
. . . . . 410 . . . . . 420 . . . . . 430 . . . . . 440 450
SIRT1 EAVRGDIF-- -------NQV VPR------- --CPRCPADE PLAIMKPEIV
SIRT3 EDIRADVM-- -------ADR VPR------- --CPVCT--- --GVVKPDIV
SIRT2 SWMKEKIF-- -------SEV TPK------- --CEDCQ--- --SLVKPDIV
SIRT6 DTVVGTMG-- -------LKA TGRLCTVAKA RGLRACR--- --GELRDTIL
SIRT7 VRVFDVTE-- -------RTA LHR---HQTG RTCHKCG--- --TQLRDTIV
SIRT4 LNPTWSAEAH GLAPDGDVFL SEEQVRSFQV PTCVQCG--- --GHLKPDVV
SIRT5 --PICPALSG KGAPEPGTQD ASIPVEKLPR CEEAGCG--- --GLLRPHVV
Consistency 215323520 0000000333 5350000010 01533*3000 0053786498
. . . . . 460 . . . . . 470 . . . . . 480 . . . . . 490 500
SIRT1 FFGE---NLP EQFHRAMKYD KDEVDLLIVI GSSLKV---R PVALIPSSIP
SIRT3 FFGE---PLP QRFLLHV-VD FPMADLLLIL GTSLEV---E PFASLTEAVR
SIRT2 FFGE---SLP ARFFSCMQSD FLKVDLLLVM GTSLQV---Q PFASLISKAP
SIRT6 DWED---SLP DRDLALADEA SRNADLSITL GTSLQI---R PSGNLPLATK
SIRT7 HFGERGTLGQ PLNWEAATEA ASRADTILCL GSSLKVLKKY PRLWCMTKPP
SIRT4 FFGD---TVN PDKVDFVHKR VKEADSLLVV GSSLQV---Y SGYRFILTAW
SIRT5 WFGE---NLD PAILEEVDRE LAHCDLCLVV GTSSVV---Y PAAMFAPQVA
Consistency 4878000464 4434326234 3346*65867 *7*7590004 8353533443
. . . . . 510 . . . . . 520 . . . . . 530 . . . . . 540 550
SIRT1 HEVPQI-LIN REPLP----- ---------- -HL-----HF DVELLGDCDV
SIRT3 SSVPRL-LIN RDLVG----- ---------- -PLAWHPRSR DVAQLGDVVH
SIRT2 LSTPRL-LIN KEKAGQSDPF LGMIMGLGGG MDFDSKKAYR DVAWLGECDQ
SIRT6 RRGGRLVIVN LQPTK----- ---------- -----HDRHA DLRIHGYVDE
SIRT7 SRRPKLYIVN LQWTP----- ---------- -----KDDWA ALKLHGKCDD
SIRT4 EKKLPIAILN IGPTR----- ---------- -----SDDLA CLKLNSRCGE
SIRT5 ARGVPVAEFN TETTP----- ---------- -----ATNRF RFHFQGPCGT
Consistency 3534581678* 4536300000 0000000000 0000033234 4644383643
. . . . . 560 . . . . . 570 . . . . . 580 . . . . . 590 600
SIRT1 IINELCHRLG GEYAKLCCNP VKLSEITEKP PRTQKELAYL SELPPTPLHV
SIRT3 GVESLVELLG WTEEMR---D LVQRETGKLD GPDK------ ----------
SIRT2 GCLALAELLG WKKELE---D LVRREHASID AQS------- ----GAGVPN
SIRT6 VMTRLMKHLG LEIPAWDGPR VLERALPPLP RPP------- ----TPKLEP
SIRT7 VMRLLMAELG LEIPAY---- --SRWQDPIF SLA------- ----TPLRAG
SIRT4 LLPLIDPC-- ---------- ---------- ---------- ----------
SIRT5 TLPEALACHE NETVS----- ---------- ---------- ----------
Consistency 3533734255 1422310001 2113211221 1110000000 0000110100
. . . . . 610 . . . . . 620 . . . . . 630 . . . . . 640 650
SIRT1 SEDSSSPERT SPPDSSVIVT LLDQAAKSND DLDVSESKGC MEEKPQEVQT
SIRT3 ---------- ---------- ---------- ---------- ----------
SIRT2 PSTSASPKKS PPPAKDEART TEREKPQ--- ---------- ----------
SIRT6 KEESPTRING SIPAGPKQEP CAQHNGSEPA SPKRERPTSP APHRPPKRVK
SIRT7 EEGSHSR--K SLCRSREEAP PGDRGAPLSS APILGGWFG- --RGCTKRTK
SIRT4 ---------- ---------- ---------- ---------- ----------
SIRT5 ---------- ---------- ---------- ---------- ----------
Consistency 1202021001 1111101001 0011011000 0000000000 0000001000
. . . . . 660 . . . . . 670 . . . . . 680 . . . . . 690 700
SIRT1 SRNVESIAEQ MENPDLKNVG SSTGEKNERT SVAGTVRKCW PNRVAKEQIS
SIRT3 ---------- ---------- ---------- ---------- ----------
SIRT2 ---------- ---------- ---------- ---------- ----------
SIRT6 AKAVPS---- ---------- ---------- ---------- ----------
SIRT7 RKKVT----- ---------- ---------- ---------- ----------
SIRT4 ---------- ---------- ---------- ---------- ----------
SIRT5 ---------- ---------- ---------- ---------- ----------
Consistency 0101000000 0000000000 0000000000 0000000000 0000000000

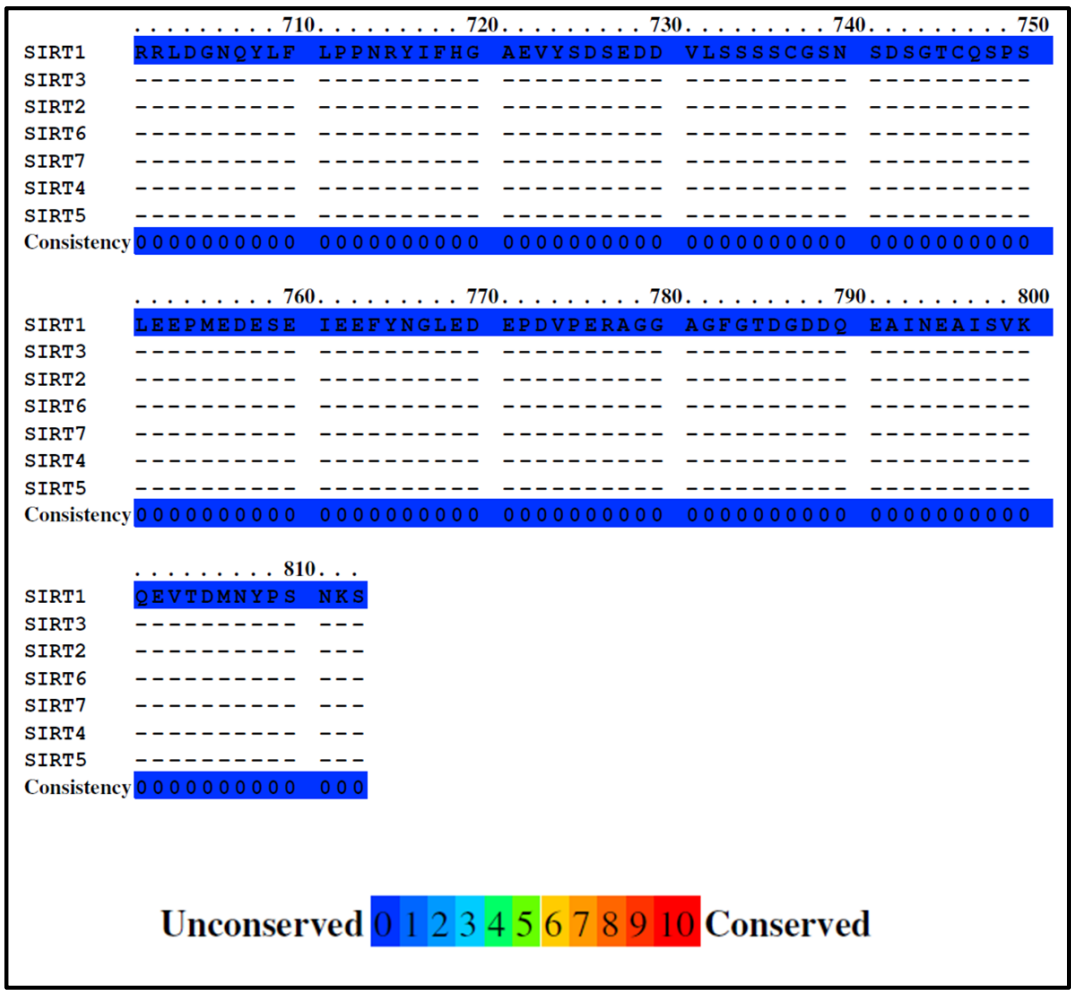


Figure 2. 8: Praline sequence analysis of SIRT1-7.

The analysis elucidates the high degree of conservation of the catalytic core consisting of ~270 amino acids. The residues are coloured by their conservation grades using the colour coding bar, with blue-through-red indicating unconserved-through-conserved residues.

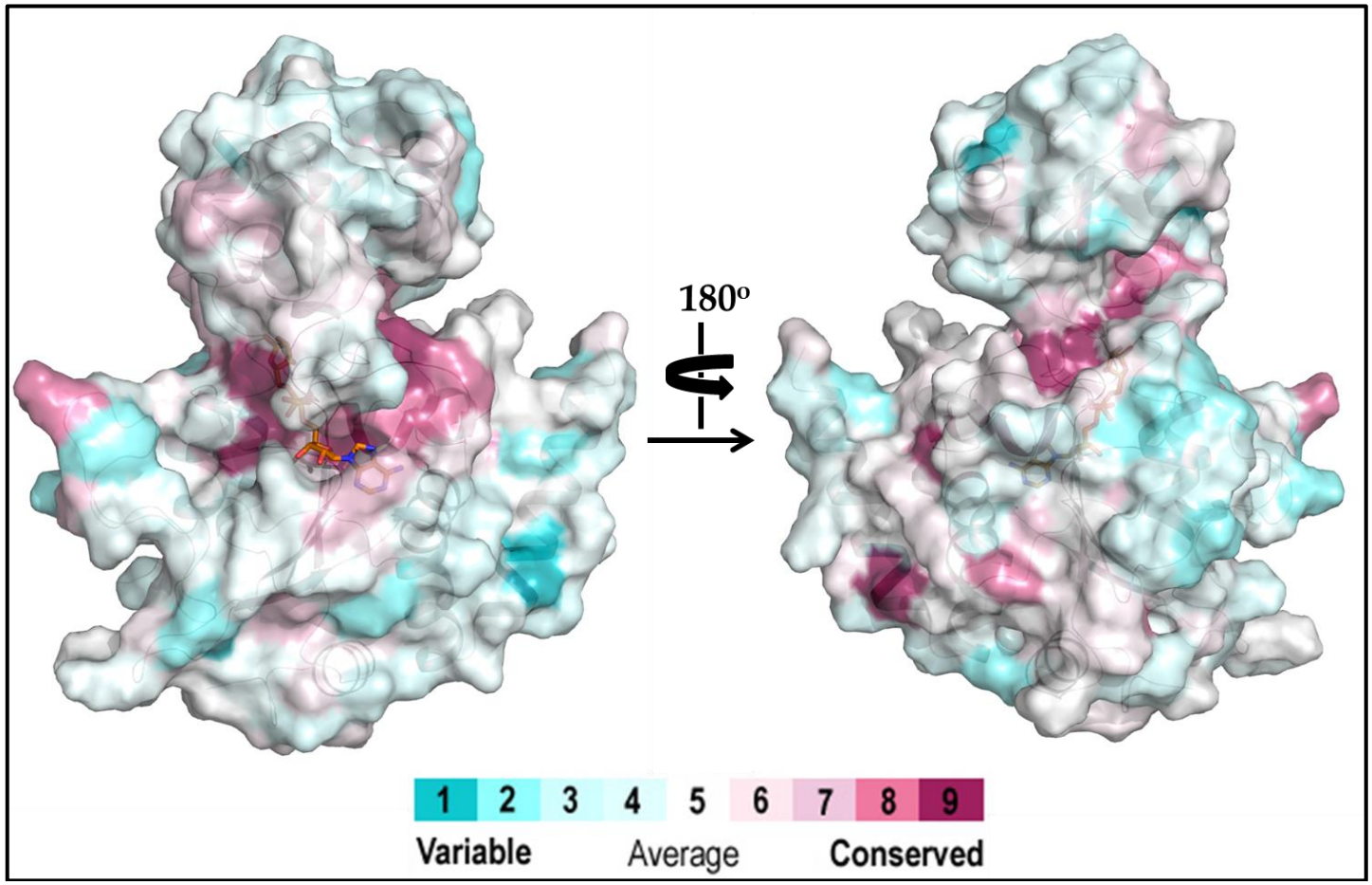


Figure 2. 9: A ConSurf analysis of SIRT2 in complex with ADPR.

(PDB: 3ZGV) ADPR is shown as sticks in orange. The amino-acids are coloured by their conservation grades using the colour-coding bar, with turquoise-through-maroon indicating variable-through-conserved. The figure reveals highly conserved substrate binding pocket among sirtuins. .

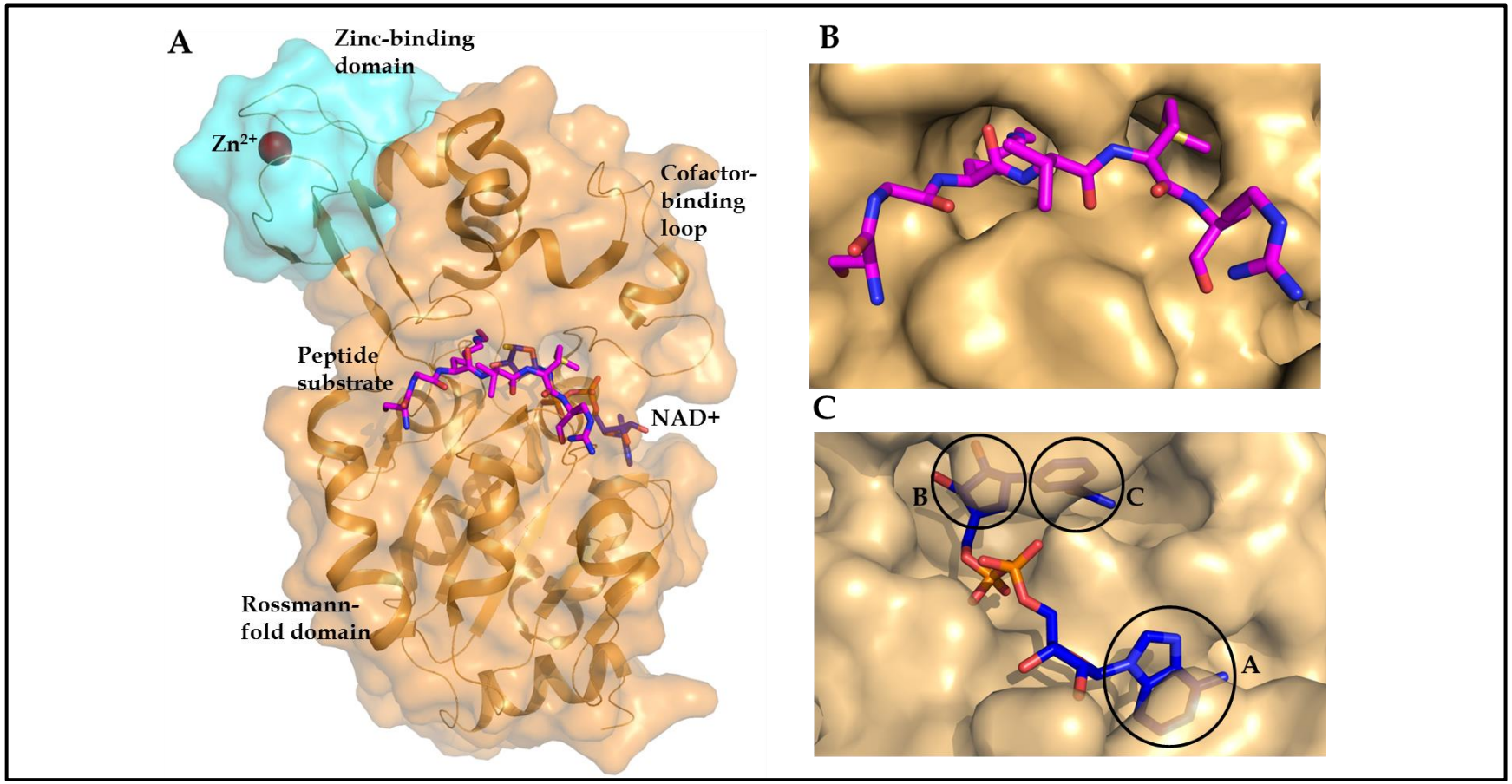


Figure 2. 10: Representative structure of a human sirtuin.

(A) Overall structure of SIRT3 bound with AceCS2 (acetylCoA synthetase 2) peptide and NAD^+ analogue (PDB: 3GLR). Key locations for sirtuin modulation are highlighted. The structure of SIRT3 is shown as cartoon model with Rossmann-fold domain and zinc-binding domain (cyan) (PDB: 3GLR). (B) The substrate binding cleft in surface representation. AceCS2 shown as pick sticks. (C) The NAD^+ binding region shown in surface representation. NAD^+ analogue shown as blue sticks (PDB: 1SZC).

2.5. Sirtuins and their implication in pathological conditions.

Sirtuins have emerged as potential therapeutic targets for the treatment of human pathologies such as metabolic, cardiovascular and other ageing-related diseases, as well as cancer.

2.5.1. Sirtuins in ageing

Sir2 and its orthologues play important roles in ageing and longevity control in various model organisms such as nematode *Caenorhabditis elegans*, fruit fly *Drosophila melanogaster*, and in mice (Kaeberlein *et al.*, 1999, Rogina and Helfand, 2004, Tissenbaum and Guarente, 2001). Studies on these organisms have also shown that Sir2 and its orthologues mediate calorie restriction–induced lifespan extension in certain genetic backgrounds. (Anderson *et al.*, 2003, Lin *et al.*, 2000, Lin *et al.*, 2002, Rogina and Helfand, 2004, Wang and Tissenbaum, 2006). Many studies have established a link between SIRT1 mediated anti-ageing effects of calorie restriction in mice (Guarente, 2013), however, SIRT1 overexpression in mice failed to support lifespan extension (Herranz *et al.*, 2010). Furthermore, their role in worm and fly lifespan has been questioned (Burnett *et al.*, 2011, Viswanathan and Guarente, 2011). Nonetheless, recently, a number of studies have reconfirmed the original hypothesis and have shown that SIRT1 does appear to promote healthy ageing by protecting against several age-related conditions (Banerjee *et al.*, 2012, Guarente, 2013, 2011, Imai and Guarente, 2014, Kanfi *et al.*, 2012, Satoh *et al.*, 2013, Schmeisser *et al.*, 2013, Viswanathan and Guarente, 2011).

The most accepted link between sirtuins and anti-ageing effects of calorie restriction is that of SIRT3 mediating the prevention of age-related hearing loss by calorie restriction (Someya *et al.*, 2010). Age-related hearing loss is characterised by the gradual loss of spiral ganglion neurons and sensory hair cells in the cochlea of the inner ear (Liu and Yan, 2007). Earlier studies revealed that calorie restriction prevents hearing loss in a variety of species and Someya *et al.* showed that SIRT3-deficient mice are resistant to the effects of calorie

restriction (Someya *et al.*, 2010). Also, SIRT3 is required for calorie restriction mediated reduction of oxidative damage in multiple tissues via regulation of the glutathione antioxidant system (Someya *et al.*, 2010). Deacetylation of manganese superoxide dismutase (MnSOD) by SIRT3 directly modulates reactive oxygen species (ROS) (Qiu *et al.*, 2010). Hence, suggesting that SIRT3 plays a major role in the regulating age-related diseases that depend on cellular levels of ROS.

The targets of the nuclear protein, SIRT1, include transcriptional proteins that are vital in energy metabolism such as PCG-1α and FOXO (forkhead box subgroup O) (Brunet *et al.*, 2004, Li *et al.*, 2007, Motta *et al.*, 2004, Rodgers *et al.*, 2005). It also regulates components of the circadian clock, such as BMAL1 (aryl hydrocarbon receptor nuclear translocator-like, a.k.a. Arntl) and period circadian protein homolog 2 (PER2) (Asher *et al.*, 2008, Nakahata *et al.*, 2008). These targets highlight the link between sirtuins, metabolism, circadian rhythm and ageing.

Recent studies on SIRT6 and longevity have revealed that SIRT6 overexpression extends lifespan in male but not female transgenic mice (Kanfi *et al.*, 2012). A significant difference was observed between male SIRT6-transgenic and male wild-type mice. The transgenic males display lower serum levels of insulin like growth factor 1 (IGF1), higher levels of IGF-binding protein 1 and altered phosphorylation levels if major components of IGF1 signalling, a key pathway in regulation of lifespan. A gender specific (males only) modest lifespan increase of 10-15% was observed (Kanfi *et al.*, 2012).

These studies provide encouraging evidence that mammalian sirtuins may play an important role as modulators of lifespan and in ageing related diseases.

2.5.2. Metabolic Diseases

As sirtuins are NAD^+ dependent enzymes, their activity is directly linked to the metabolic state of the cell. All the members of the sirtuin family have been shown to play a role in regulating metabolism through their involvement in

metabolic pathways and energy homeostasis [Table 2.4]. Sirtuins often play a role in helping the cell adapt to periods of low energy input.

In addition to glucose and lipid metabolism (discussed below), sirtuins have important functions in additional metabolic pathways. SIRT3 upregulates the electron transport chain via succinate dehydrogenase (SDHA) (Finley *et al.*, 2011) and the urea cycle via ornithine transcarbamoylase (OTC) (Hallows *et al.*, 2011). It also activates ketone body synthesis via HMGCS2 (3-hydroxy-3-methylglutaryl-CoA synthase 2) (Shimazu *et al.*, 2010). SIRT5 regulates the urea cycle through demalonylation and desuccinylation of carbamoyl phosphate synthetase 1 (CPS1) (Du *et al.*, 2011).

Table 2. 4: The role of sirtuins in regulating metabolism in various tissues.

Tissue	Sirtuin	Function
Brain	SIRT1	Decrease weight gain on high calorie diet intake.
Liver	SIRT1	Decrease lipogenesis (SREBP-1c) Increase gluconeogenesis (PCG1α)
	SIRT2	Increase gluconeogenesis (PEPCK)
	SIRT3	Increase fatty acid oxidation (LCAD) Increase ketone body production (HMGCS2) Increase urea cycle (OTC)
	SIRT5	Increase urea cycle (CPS1)
	SIRT6	Increase fatty acid oxidation
Pancreas	SIRT1	Increase glucose tolerance Increase insulin secretion
	SIRT4	Decrease insulin secretion
Adipose tissue	SIRT1	Increase lipolysis (FOXO1)
Skeletal muscle	SIRT1	Increase fatty acid oxidation
	SIRT3	Increase fatty acid oxidation(LCAD)
	SIRT6	Decrease glucose uptake(H3K9Ac)

Glucose and glutamine metabolism

Gluconeogenesis is the metabolic process by which organisms produce glucose from non-carbohydrate carbon precursors such as pyruvate and lactate. Glucose, the primary nutrient for cell survival and proliferation, is tightly regulated throughout tissues. The liver, muscle and the pancreas act as modulators of glucose homeostasis. In the cytoplasm of the cell, glucose gets

converted to pyruvate through glycolysis. Pyruvate then enters the citric acid cycle (or tricarboxylic acid cycle or TCA cycle in the mitochondria) to generate NADH. This molecule is in turn used by the oxidative phosphorylation pathway to generate adenosine triphosphate (ATP). 30 mols of ATP per mol of glucose is generated in this process. However, in certain situations, pyruvate gets diverted in the cell to produce lactate which is a less efficient way to produce ATP but a crucial method for ATP production in cases of impairment of oxidative phosphorylation (*e.g.* hypoxia).

SIRT1 has been reported to modulate both gluconeogenesis and glycolysis by regulating important metabolic factors such as PGC1α and FOXO (Canto and Auwerx, 2012).

SIRT2 has shown to promote gluconeogenesis through deacetylation and stabilisation of the rate limiting enzyme phosphoenolpyruvate carboxykinase (PEPCK), thus participating in glucose metabolism (Jiang *et al.*, 2011).

Mitochondrial SIRT3, regulates metabolic proteins such as isocitrate dehydrogenase 2 (IDH2) which plays an important role in the TCA cycle (Someya *et al.*, 2010). SIRT3 also protects the mitochondria from oxidative stress by deacetylating FOXO3a which in turn increases the expression of antioxidant genes (Tseng *et al.*, 2013).

Glutamine is the main source of replenishment in the TCA cycle of proliferating cells as it is a valuble source of α-ketoglutarate (α-KG). SIRT4 inhibits glutamine entry to the TCA cycle under genotoxic stress and prevents dysregulated proliferation and genomic instability (Jeong *et al.*, 2013).

In recent years, the role of SIRT6 as a critical modulator of glucose homeostasis has come to light. SIRT6 knockout mice die earlier in life compared to the wild-type control due to fatal hypoglycemia. These knockout mice exhibited increases expression of the glucose transporter GLUT1, leading to increased glucose uptake in skeletal and brown adipose tissue (Mostoslavsky *et al.*, 2006, Zhong *et al.*, 2010). SIRT6 negatively regulates HIF-1α-dependent transcription

by deacetylating H3K9Ac at the promoter of metabolic genes such as GLUT1, lactate dehydrogenase A (LDHA), PDH kinase 1 (PDHK1). This suggests that the increased glucose uptake seen in SIRT6 knockout mice can be reversed by the treatment with HIF-1α inhibitor (Zhong *et al.*, 2010).

Lipid metabolism

Lipids are an important constituent of the cellular membrane and is an important source of energy in normal cell physiology. Their synthesis, storage and expenditure are tightly regulated by various physiological requirements of the cell such as fasting and nutrient availability. Lipid synthesis takes place in the liver in order to store energy inside white adipose tissue. Fatty acid synthesis occurs in the cytoplasm and fatty acid oxidation to produce acetyl-CoA takes place in the mitochondrial matrix. Acetyl-CoA is a key molecule of the TCA cycle and is used to produce ATP. Depending on the subcellular localisation of sirtuins, they actively modulate fatty acid synthesis (cytoplasm) and oxidation (mitochondria) via transcriptional or posttranslational regulation by acting as energy and redox sensors (Lomb *et al.*, 2010).

Lipid catabolism is important during fasting. Under situations of calorie restriction, SIRT1 stimulates lipid breakdown through the transcription factor FOXO1 which in turn induces the expression of adipose triglyceride lipase (ATGL), a rate-limiting lipolytic enzyme (Chakrabarti *et al.*, 2011). SIRT1 targets triglyceride synthesis in the liver by deacetylating and suppressing sterol-response element-binding protein 1c (SREBP-1c)-dependent transcription (Ponugoti *et al.*, 2010, Walker *et al.*, 2010). SIRT1 overexpression was shown to protect mice placed on a high-fat diet from hepatic steatosis (fatty liver) (Pfluger *et al.*, 2008) in one study, while opposite results were observed in another (Qiang *et al.*, 2011). These studies attributed the different outcomes to SIRT1 induced inhibition of Creb, a transcription factor which activates fatty acid metabolism and gluconeogenesis. Like SIRT1, SIRT2 deacetylates FOXO1 and inhibits adipocyte differentiation (Wang and Tong, 2009).

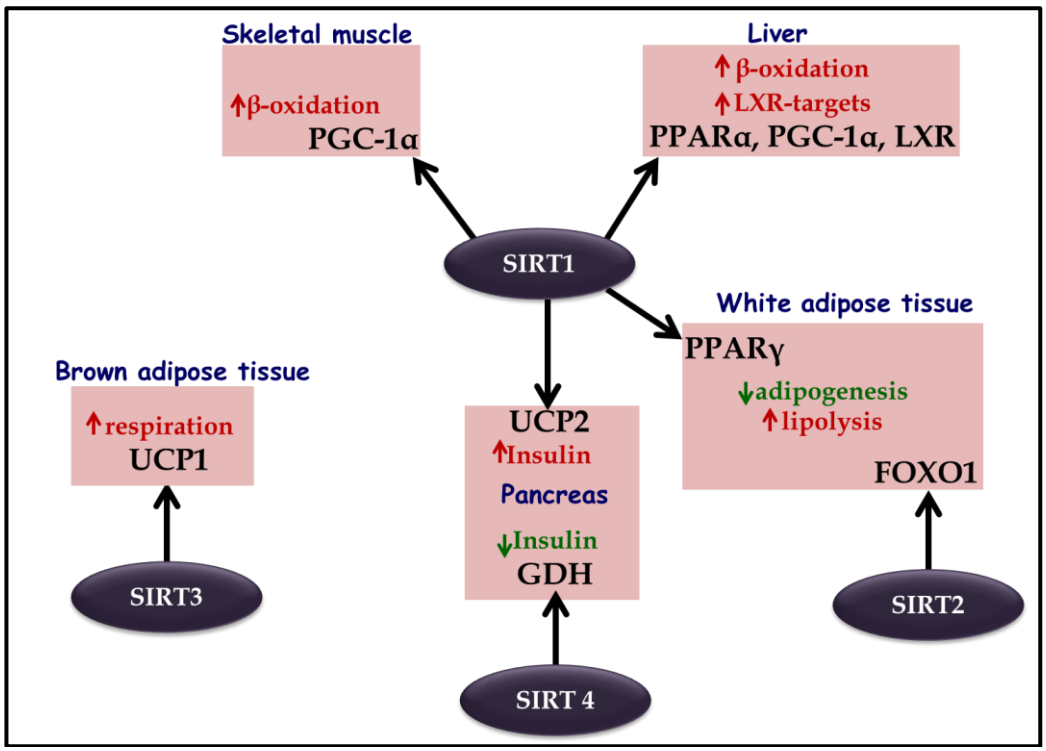


Figure 2. 11: Sirtuin involvement in lipid metabolism.

In white adipose tissue, SIRT1 and SIRT2 repress adipogenesis and promote lipolysis by inhibiting PPARγ and FOXO1, respectively. In the liver, SIRT1 promotes β-oxidation by stimulating PPARα/PGC-1α. SIRT1 also deacetylates LXR thereby inducing expression of LXR target genes. In the skeletal muscle, SIRT1 promotes β-oxidation through PGC-1α. SIRT1 and SIRT4 modulate pancreatic insulin secretion. In the brown adipose tissue, SIRT3 promotes increased mitochondrial respiration by inducing UCP1.

Metabolic SIRT3 and SIRT4 play key roles in fatty acid oxidation. SIRT3 expression is upregulated during fasting. Genetic ablation of SIRT3 affects various enzymes, including residues K318 and K322 of long-chain acyl-CoA dehydrogenase (LCAD), resulting in a decrease in fatty acid oxidation in liver mitochondria (Bharathi *et al.*, 2013, Hirschey *et al.*, 2010). SIRT4 is a repressor of malonyl-CoA decarboxylase (MCD), a core enzyme that balances malonyl-CoA and acetyl-CoA in mitochondria and modulates lipid anabolism and catabolism. SIRT4 deacetylates and inhibits MCD activity, favouring fatty acid synthesis over oxidation (Laurent *et al.*, 2013).

SIRT6 along with SIRT1 promotes fatty acid oxidation. SIRT1 and the transcription factors FOXO3a and nuclear respiratory factor 1 (NRF1) form a complex on the SIRT6 promoter and promote fatty acid oxidation in wild-type mice (Kim *et al.*, 2010b). A liver-specific deletion of SIRT6 in mice casues hepatic steatosis due to low levels of fatty acid oxidation while transgenic overexpression of SIRT6 protects against accumulation of visceral fat when the mice are placed on a high-fat diet (Kanfi *et al.*, 2010).

In a recent study SIRT7, like other members of the sirtuin family, was revealed to be involved in lipid metabolism. SIRT7 was shown to alleviate high-fat diet-induced hepatic steatosis by co-repressing Myc transcriptional activity and decreasing endoplasmic reticulum stress in liver (Shin *et al.*, 2013).

2.5.3. Cancer

Detecting and responding to life-threatening environmental changes induced by metabolic fluctuation and stress remains one of the greatest challenges facing organisms. These conditions are linked to the onset of many pathological conditions. Sirtuins participate in an elaborate network of interactions and their targets that are involved in cell growth, cell cycle progression, genomic integrity and cell death. This enables sirtuins to play vital roles as cell sensors, modulators and responders to stress. Stress and environmental conditions can induce mutations that lead to cell transformation and cancer initiation. This is coupled with metabolic reprogramming to compensate for the energy

requirement for tumour cell progression. The ability of sirtuins to regulate genomic stability and other cellular functions makes them ideal candidates for the control of tumourigenesis. In cancer, sirtuins have been reported to have a dual role. While some sirtuins prevents organisms from cancer by helping to protect DNA from damage and oxidative stress, maintaining genomic stability and limiting replicative life span, others promote cell survival under stress conditions and promote tumourigenesis. As sirtuins are key modulators in various pathways, understanding their function in cancer would unearth new targets and tools for the control of tumour growth.

Sirtuins play a complex role in both promoting and/or supressing tumourigenesis. The functional loss of sirtuin genes, which are involved in maintaining genome integrity and DNA repair, promotes tumourigenesis due to genomic instability upon their loss (Roth and Chen, 2014). This is contradicted by the requirement of sirtuins for these same processes to allow cancerous cells to survive, proliferate, repair genomic events and evolve. This dual role of sirtuins in cancer has been a topic of discussion and has eluded the understanding of its role in cancer.

p53 is vital in the regulation of the cell cycle and functions as a tumour suppressor to prevent cancer. In normal cell physiology, p53 levels are low and increase when the cell is subjected to stress. Its major functions include growth arrest, DNA repair and apoptosis. Unusually high levels of p53 accelerate ageing (an established hallmark of cancer) due to excessive apoptosis. SIRT1 and SIRT7 deacetylate p53 and reduce its influence on cell cycle control during stress and in response to DNA damage. Overexpression of SIRT1 deactivates p53 which results in a reduction of the cell's ability to respond to stress and DNA damage (Luo *et al.*, 2001, Vaziri *et al.*, 2001). Along with p53, SIRT1 overexpression can also block FOXO, E2F1, retinoblastoma protein (Rb), B-cell lymphoma 6(BCL6) and Ku70 via deacetylation (Bosch-Presegue and Vaquero, 2011, Heltweg *et al.*, 2006, Wang *et al.*, 2006). Increased levels of SIRT1 have been associated with various cancers including prostate (Huffman *et al.*, 2007), acute myeloid leukaemia (AML) (Bradbury *et al.*, 2005), primary colon (Stunkel *et al.*,

2007) and other non-malignant skin cancers (Hida *et al.*, 2007). Inhibition of tumour suppressors by SIRT1 along with its upregulation in various cancers, provides evidence supporting the role of SIRT1 in tumour initiation and progression via blocking senescence and apoptosis.

The tumour suppressive function of SIRT1 in cancer is corroborated by its role in maintaining genome stability through chromatin regulation and DNA repair (Bosch-Presegue and Vaquero, 2011, Saunders and Verdin, 2007, Stunkel *et al.*, 2007). Decreased SIRT1 levels have been reported in glioblastoma, bladder carcinoma, male germ cell tumour prostate carcinoma and ovarian cancers (Wang *et al.*, 2008). The role of SIRT1 as a tumour suppressor has been substantiated by other studies (Ford *et al.*, 2005, Kojima *et al.*, 2008, Pruitt *et al.*, 2006).

Akin to SIRT1, SIRT2 has both pro-and anti-cancer roles. Control of cell-cycle progression by SIRT2 is essential for prevention of tumours as it is suppressed in gliomas (Hiratsuka *et al.*, 2003) and melanomas (Lennerz *et al.*, 2005). SIRT2 is down-regulated in some cancers and its loss is shown to compromise the mitotic checkpoint, contributing to genomic instability and tumourigenesis (Dryden *et al.*, 2003, Inoue *et al.*, 2007). On the other hand, SIRT2 can promote oncogenic phenotypes. SIRT2 is increased in AML *in vitro* (Dan *et al.*, 2012). Knockdown of SIRT2 is shown to induce p53 accumulation and promotes apoptosis of cancer cells (Li *et al.*, 2011, Peck *et al.*, 2010).

Analogous to the bifurcated roles of SIRT1 and SIRT2 in cancer, SIRT3 has roles in tumour suppression and progression. Under normal conditions, it has been shown to promote apoptosis while under stress conditions it promotes cell survival (Allison and Milner, 2007, Kim *et al.*, 2010a).

More recently, SIRT5 overexpression has been implicated in a study of pancreatic cancers (Ouaissi *et al.*, 2008). The tumour suppressive role of SIRT6 is supported by its role as a guardian of genome stability (Mahlknecht *et al.*, 2006).

Elevated levels of SIRT7 in human breast, thyroid and liver cancers have been reported (Ashraf *et al.*, 2006, de Nigris *et al.*, 2002, Kim *et al.*, 2013b). SIRT7 is a lysine deacetylase that selectively removes a histone mark, acetylated H3K18 (H3K18Ac). Deacetylation of H3K18 (which is hypoacetylated in many cancers including prostate cancer) by SIRT7 is necessary for maintaining essential features of human cancer cells, including anchorage-independent growth and escape from contact inhibition (Barber *et al.*, 2012, Seligson *et al.*, 2009). These studies direct at SIRT7 providing a growth advantage to cancer cells.

Further studies are needed to delineate the precise roles of sirtuins in cancer. The bifurcated role of sirtuins in cancer can be accredited in part to the nature of the genes involved in the cell's genome maintenance systems (Roth and Chen, 2014). Deciphering sirtuin functions would have a significant contribution to designing precise modulations of selective sirtuins to help cancer prevention and treatment.

2.5.4. Neurodegenerative diseases

Sirtuins play an important role in maintaining neuronal health during ageing and have been shown to be beneficial in multiple models of neuropathology. During neuronal development, SIRT1 promotes axonal elongation, neurite outgrowth and dendritic branching. It also plays a role in memory formation by modulating synaptic plasticity. It regulates the hypothalamic functions that control feeding behaviour, endocrine function and circadian rhythmicity (Herskovits and Guarente, 2014).

The protective role of SIRT1 against neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotropic lateral sclerosis (ALS), spinal and bulbar muscular atropy (SBMA), prion disease, multiple sclerosis and psychiatric disorders have been reported (Herskovits and Guarente, 2014). Brain-specific SIRT1 knockout in mice increases β-amyloid plaque formation, a hallmark of Alzheimer's disease. Overexpression of SIRT1 reverses this effect (Donmez *et al.*, 2010). This may be due to activation of transcription of the *ADAM-10* gene that encodes for α-

secretase which helps the brain avoid plaque formation. Another hallmark of Alzheimer's is tau proteins which were shown to be destabilized by SIRT1 deacetylation, thereby preventing protein aggregation (Min *et al.*, 2010). Similarly, SIRT1 activates molecular chaperones such as HSP70 by deacetylation of HSF1 in response to α-synuclein aggregation-induced stress. α-synuclein is a key molecule in the pathogenesis of synucleinopathy including dementia with Lewy bodies, Parkinson's disease and multiple system atrophy (Donmez *et al.*, 2012). In Huntington's disease, SIRT1 overexpression improves motor function in mice while brain-specific knock-out worsens disease-related pathology (Jeong *et al.*, 2012, Jiang *et al.*, 2012). In contrast, SIRT2 inhibition has shown to lower toxicity in a striatal neuron model of Huntington's disease (Luthi-Carter *et al.*, 2010).

Little is known about the role of other sirtuins in the brain. Further understanding of these roles and rational drug design that modulate specific sirtuins may offer a promising approach to treat neurodegenerative disorders.

2.5.5. Cardiovascular diseases

Cardiovascular disease (CVD) is among the leading causes of death globally. It is marked by the deterioration of heart and blood vessel function. As calorie restriction improves both endothelial and heart function (Weiss and Fontana, 2011), the functions of sirtuins in CVD have been explored. The role of SIRT1 in endothelial function, vessel inflammation, vascularisation and cholesterol metabolism has been well established (Haigis and Sinclair, 2010). SIRT1, SIRT3, SIRT6 and SIRT7 are key modulators of cardiac hypertrophy.

Deacetylation and regulation of p53 by SIRT1 and SIRT7 enables them to protect against cardiac hypertrophy (Pillai *et al.*, 2005, Vakhrusheva *et al.*, 2008). SIRT1 also binds to PPARα favouring the deacetylation of PGC-1α and prevents downregulation of fatty acid oxidation genes (Planavila *et al.*, 2011).

The mitochondrial SIRT3 is an endogenous negative regulator of cardiac hypertrophy which protects the heart by suppressing cellular levels of ROS and

deactivation of protein kinase B (AKT). SIRT3 dependent deacetylation of FOXO3a leads to its nuclear localization and enhances the expression of manganese superoxide dismutase (MnSOD) which reduces ROS levels (Sundaresan *et al.*, 2009). Moreover, SIRT3 deacetylates and activates liver kinase B1 (LKB1) thus augmenting the activity of the LKB1-AMPK pathway and supresses AKT phosphorylation (Pillai *et al.*, 2010).

SIRT6 directly attenuates IGF/AKT signalling by binding to and suppressing the promoter of IGF signalling-related genes by interacting with c-Jun and deacetylating H3K9 (Sundaresan *et al.*, 2012). Increased levels of SIRT6 suppress NF-κB activation and protect cardiomyocytes from hypertropic responses *in vitro* (Yu *et al.*, 2013).

2.5.6. Inflammation

Inflammation is one of the hallmarks of many diseases including cancer and cardiac diseases. The anti-inflammatory role of sirtuins, especially SIRT1 and SIRT3, has been brought to light by recent research. SIRT1 exhibits pronounced anti-inflammatory properties. The endotoxin, lipopolysaccharide (LPS) is a predominant microbial mediator that activates the pro-inflammatory cascade in the human digestive tract and plays a major role in various inflammatory-related diseases (Alexander and Rietschel, 2001, Miller *et al.*, 2005). Increased levels of LPS-stimulated phosphorylation of JNK and c-JUN along with enhanced expression of inflammatory cytokines such as TNF-α, IL-1β and IL-6 in SIRT1 knockdown mice macrophage were observed (Yoshizaki *et al.*, 2009). Lung inflammation following exposure to airborne particulate matter may also be prevented by SIRT1 (Wu *et al.*, 2012). Deacetylation of residues K310 of the NF-κB subunit p65 by SIRT2 results in downregulation of the immune response (Rothgiesser *et al.*, 2010). SIRT7 knockout mice had elevated myocardial levels of several cytokines demonstrating the role of SIRT7 in inflammation (Vakhrusheva *et al.*, 2008).

2.6. Therapeutic potential

The role of sirtuins in cell survival, fatty acid metabolism, glucose homeostasis, genomic stability and oxidative stress makes them potential pharmaceutical targets. Activators and inhibitors to treat diseases, regulate metabolism and extend lifespan have been extensively researched.

Among members of the sirtuin family, SIRT1 targeting molecule development has been in full swing since its early discovery in a variety of human disorders such as metabolic diseases, cancer and ageing related diseases. Drug discovery work on other sirtuins has been less explored. Inhibition of SIRT1 could be useful for cancer treatment while activators have potential use in life extension. Over the past decade, a variety of small-molecule SIRT1-activating compounds (STACs) or inhibitors have been published [Figure 2.12]. Of these only nicotinamide (NAM) is a physiological inhibitor.

Activators of SIRT1 can be used to treat a variety of conditions such as inflammation, neurodegeneration, regulation of glucose metabolism by regulating insulin secretion, regulating ageing process and beneficial metabolic effects in calorie restriction (as discussed in the Section 2.5). Activators of SIRT1 can mimic the effects of calorie restriction which provides a new weapon in the treatment of obesity and diabetes (Haigis and Sinclair, 2010). Polyphenols such as resveratrol and other small molecule compounds have been developed to increase the deacetylase activity of SIRT1 (Alcain and Villalba, 2009a, Howitz *et al.*, 2003, Milne *et al.*, 2007). Resveratrol is a pleiotropic molecule and has shown to activate the AMPK pathway resulting increased cellular NAD^+ concentration and subsequent stimulation of SIRT1 activity (Baur *et al.*, 2006, Park *et al.*, 2012). However, the exact mechanism of SIRT1 activators remains unclear.

Increased SIRT1 levels in neurodegenerative diseases and many cancers including prostate, AML and colon cancer threw light on the importance of developing inhibitors of sirtuins (as discussed in the Section 2.5.3) (Alcain and Villalba, 2009b). In recent years a number of sirtuin inhibitors have been discovered and characterised. In addition to nicotinamide, the physiologic

sirtuin inhibitor, specific inhibitors like splitomycin and its analogues, AGK2, tenovin, surfactin, sirtinol, suramin and cambinol have been developed (Bedalov *et al.*, 2001, Heltweg *et al.*, 2006, Lain *et al.*, 2008, Sanders *et al.*, 2010, Trapp *et al.*, 2006).

Sirtuins have grown as potential therapeutic targets for age-related diseases and metabolic dysfunction. Conflicting data supporting the bifurcated role of sirtuins means that highly specific modulation of sirtuin activity is required. Altering its activity, despite any health benefits, could result in increased health risks and compromise safety. Given the complex role of each sirtuin, future effort is needed to develop more selective sirtuin modulators for human use to treat cancer and other diseases.

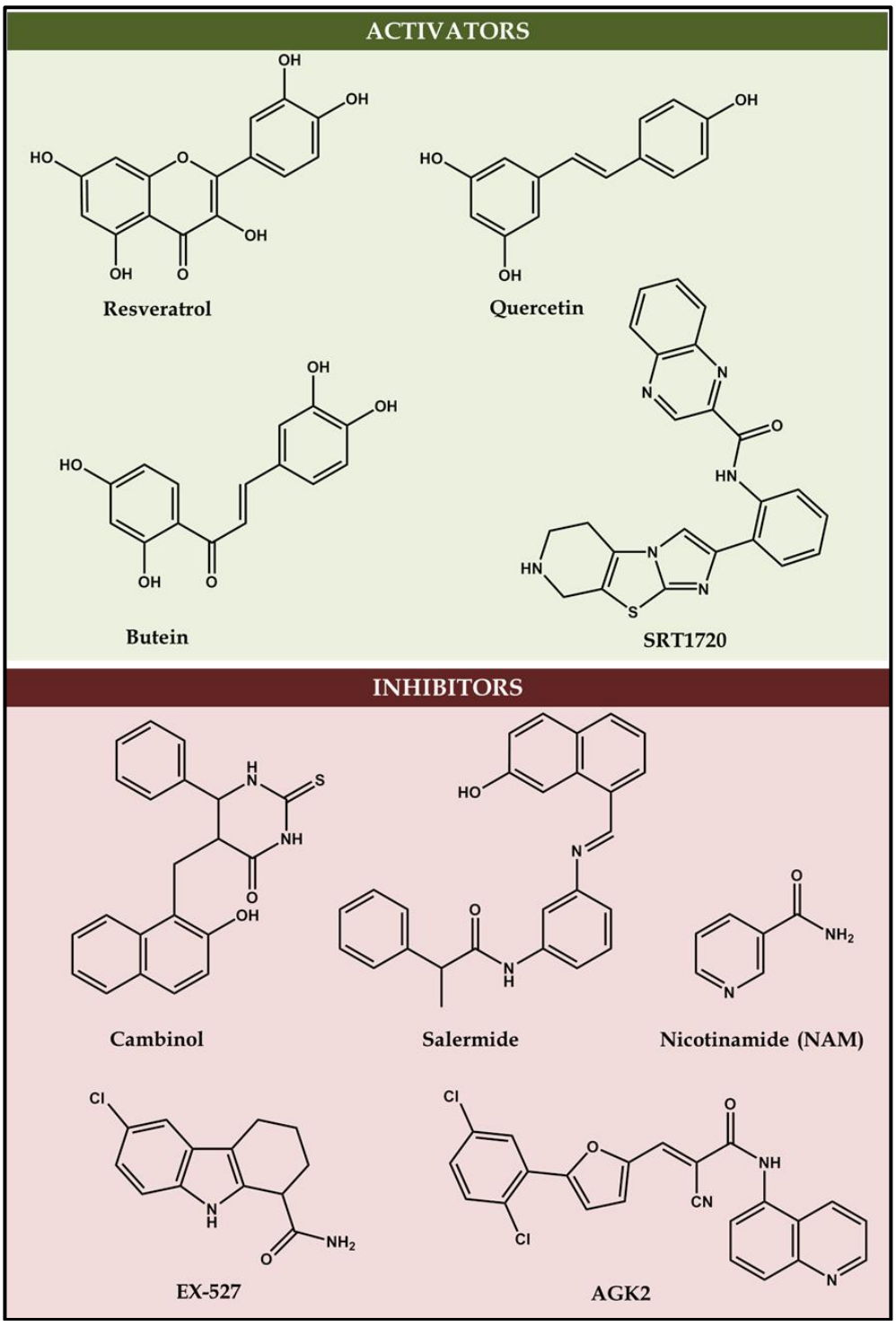


Figure 2. 12: Chemical activators and inhibitors of sirtuins.

2.7. Aims and Objectives

Sirtuins are an ancient family of proteins found in all life forms. They display a high degree of structural and functional conservation from bacteria to humans. The founding member of this vital class of histone deacetylases, yeast Sir2 (*s*ilent mating type *i*nformation *r*egulator 2), was identified as a gene required for maintaining silent chromatin (Shore *et al.*, 1984). In a NAD^+-dependent manner, sirtuins perform deacylation from lysine-modified ε-amino groups or NAD^+-dependent protein ADP-ribosylation. The presence and enzymatic activity of their homologues in other species piqued curiosity (Tanner *et al.*, 2000). Extensive work to deepen our understanding of the biological function and substrates of sirtuins was undertaken. Reports linking sirtuins to longevity and their role in ageing and age-related diseases further provoked general interest (Houtkooper *et al.*, 2012). Their unique enzymatic functions and potential therapeutic role in ageing and health issues including cancer have garnered the interest of researchers to unravel the molecular role of sirtuins in various physiological pathways. As this area of molecular science consolidates and advances, the sirtuin family of proteins is gaining significance in human biology and disease. Although a myriad of questions regarding the role of sirtuins remain unanswered, there is a growing body of evidence indicating that this research should be prioritised.

Sirtuins show a strong potential to become valuable predictive and prognostic markers for disease and as therapeutic targets for the management of a variety of cancer types and age-related diseases. However, due to the interplay of sirtuins, a detailed molecular understanding of each sirtuin is essential for the development of selective sirtuin modulators. This can be explored using a powerful structural biology technique such as X-ray crystallography which will help in understanding the mechanism of sirtuin activity at the molecular level.

The crystal structures of some of the human sirtuins have been deduced but growing evidence suggests a unique role for each sirtuin and highlights the importance of elucidating their native structures and structures in complex with various substrates. Their ability to recognise a wide variety of substrate

sequences indicates a high degree of adaptability of this class of enzymes to the substrates. It is thus essential to obtain the structures of different isoforms and enzyme states including substrate bound forms for obtaining a reliable picture of the dynamics and preferred conformations of their binding pockets for drug development.

To answer specific questions such as the difference in the deacetylation and ribosylation activity among mammalian sirtuins, the biological role of regions N- and C-terminal to the catalytic core and inhibition mechanism of synthetic sirtuin inhibitors the structure determination of sirtuin complexes containing reaction intermediates at various stages of the reaction is required.

This study of sirtuins focuses on understanding the structural biology of SIRT1, SIRT2 and SIRT7. Structural studies of SIRT7 would help understand the activity of this enzyme and substrate binding. The crystal structure of SIRT2 in its native form (Finnin *et al.*, 2001) and in complex with ADPR (Moniot *et al.*, 2013) as well as in complex with a macrocyclic peptide S2iL5 (Yamagata *et al.*, 2014) have been elucidated. Crystal structures of SIRT2 in complex with its inhibitors and NAD^+ analogues would aid in selective modulation of SIRT2 activity. SIRT1 being the largest member of the sirtuin family was selected for structural studies to identify the differences in its structures contributing to the high degree of deacetylase activity.

In order to perform detailed biochemical and structural analyses, pure protein is required. The production of pure protein and its subsequent crystallisation remain bottlenecks in this field of structural biology. For this study, cloning of the gene of interest into appropriate expression vectors and its transformation into carefully selected expression cell lines were initially performed. The purified proteins, alone and in complex with various substrates and inhibitors, were used for crystallisation screening and optimisations.

Chapter 3. MATERIALS AND METHODS

This chapter summarises the experimental procedures used throughout this thesis. Further details of each method are given in the relevant chapters.

3.1. Chemicals and reagents

All chemicals and reagents used were purchased from Sigma-Aldrich, Kent, U.K. unless otherwise mentioned.

3.2. Bacterial strains and plasmids

Several vectors and cell lines were employed for the cloning and expression of hGal-7 and sirtuins. The expression plasmids were all carefully selected based on features such as compatibility, the incorporation of purification/fusion tags, the presence of cleavage sites, and antibiotic resistance. A range of different *E. coli* cell lines were chosen as hosts for protein expression. The advantages of fast growth at a high density in an inexpensive medium, as well as the ease of genetic manipulation make *E. coli* a valuable tool in recombinant protein expression. The bacterial strains and plasmids used in this thesis are highlighted below.

Competent cells

E. coli One Shot TOP 10 chemically competent cells (Invitrogen, CA, USA) were used for cloning and propagation.

Expression strains

Several strains of *E. coli* were used for protein expression which included BL21-CodonPlus (DE3)-RIPL competent cells (initially purchased from Stratagene, CA, USA) prepared in house, Rosetta-gammi™ 2(DE3)pLysS competent cells (Novagen, Darmstadt, Germany) and ArcticExpress (DE3) RIL competent cells (Agilent, CA, USA).

Expression plasmids

Ligation independent cloning (LIC), shown in Figure 3.1, was developed in the 1990's as an alternative to the more traditional restriction method of

enzyme/ligase cloning. The number of vector choices, ease of cloning and low cost of this method often makes it favourable compared to other techniques. Many of the commercially available LIC vectors have been designed to include a protease cleavage site, most commonly enterokinase (Ek) or Factor Xa (Xa). The Ek/LIC method was employed where possible in this thesis, although restriction enzyme/ligase cloning was also used. A summary of the vectors used is given in table 3.1, while further details on construct design can be found in the relevant results chapter. All vectors were purchased from Novagen, Darmstadt, Germany except pGEx-6p-1 which was purchased from GE Healthcare, U.K.

The vector maps for the expression plasmids used are provided in appendix A.

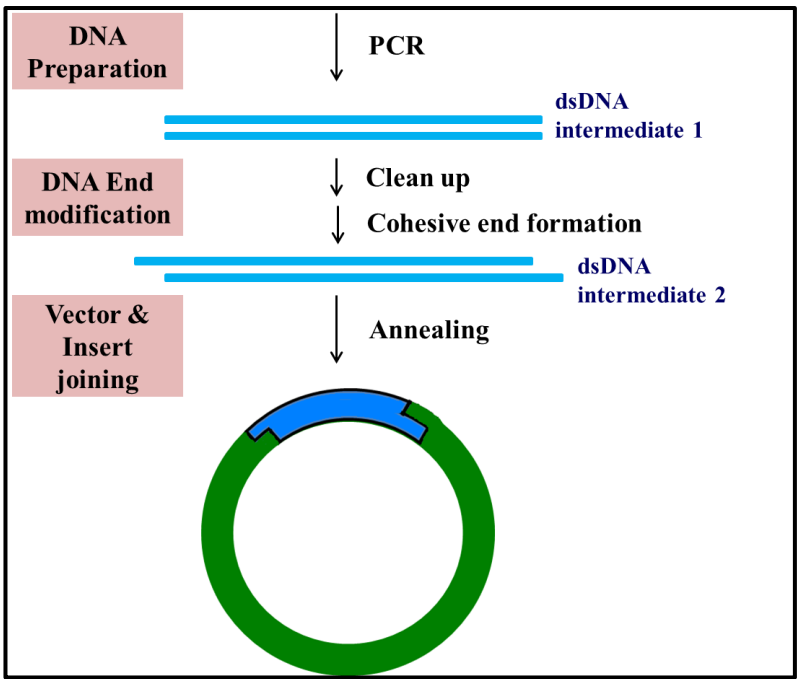


Figure 3. 1: Ligation independent cloning (LIC).
Inserts are PCR amplified and vectors are made liner either by restriction digestion or by PCR. T4 polymerase is then used to create overhangs with complementarity between the vector and insert. This technique allows the efficient creation of seamless recombinant plasmids.

Table 3. 1: Expression plasmids used for hGal-7 and sirtuin expression.

Vector	Promoter	Antibiotic Resistance	Tags and Fusion Partners	Protease Cleavage Site	Plasmid Type
pET-22b	T7-lac	Ampicillin	Signal Sequence C-terminal His-Tag	None	Bacterial Expression
pGEX-6P-1	tac	Ampicillin	N-terminal GST-Tag	Rhinovirus 3C protease (PreScission protease)	Bacterial Expression
pTriEx-4 Ek/LIC	T7-lac (*E. coli*), p10 (Insect) and CMVie (Mammalian)	Ampicillin	N-terminal His-Tag N-terminal S-Tag C-terminal His-Tag (optional) C-terminal HSV-Tag (optional)	Enterokinase Thrombin	Bacterial, Insect and Mammalian Expression
pET-46 Ek/LIC	T7-lac	Ampicillin	N-terminal His-Tag N-terminal S-Tag C-terminal His-Tag	Enterokinase Thrombin	Bacterial Expression
pET-41 Ek/LIC	T7-lac	Kanamycin	N-terminal His-Tag C-terminal S-Tag	Enterokinase	Bacterial Expression

3.3. Media and supplements

Several different types of media were used for bacterial growth depending on the requirements of the specific cell line. For example, the competent cells used in subcloning procedures were grown in Miller Luria Bertani (LB) medium (Sigma-Aldrich Co., Kent, U.K.). For the growth of bacterial colonies in petri dishes LB media was supplemented with Agar (15 g/L).

For protein expression, bacterial cell growth was tested in a wide range of media including Luria Bertani (LB), Terrific Broth (TB) and Auto-Induction (AI). Additional specialised media including modified TB, HIS, ZXY7, SOC media, glucose M9Y, superbroth, TYGPN and NZCYM media were also employed. The composition for each medium is given in appendix B.

Prior to inoculation all media was autoclaved at 120 ^{o}C for 15 min and supplemented with the appropriate antibiotics. Stock solutions of ampicillin (100 mg/mL, chloramphenicol (34 mg/mL), gentamycin (20 mg/mL) and kanamycin (30 mg/mL) were prepared and stored at -20 °C, such that a 1/1000 dilution could be used when required.

3.4. Molecular biology

Introduction of a foreign DNA molecule into a replicating cell permits the cloning or amplification (production of many identical copies) of that DNA. To clone a nucleotide sequence of interest, the cellular DNA is cleaved with a specific restriction enzyme and the cleaved DNA fragment is joined to a cloning vector to form a recombinant DNA plasmid. The plasmid is then transformed into a single host cell where it is replicated. The process of introducing foreign DNA into a cell is called transformation (bacteria and yeast) or transfection (higher eukaryotes). The multiplication of the host cell with each new cell carrying the recombinant plasmid is called cloning. The cloned DNA is eventually released from its vector by cleavage (using an appropriate *restriction endonuclease*) and is isolated. This fragment is then ligated to an expression

vector and transformed into an expression host cell. This section gives details of the methods used for cloning and transformation of the gene of interest.

3.4.1. Agarose gel electrophoresis

Agarose gel electrophoresis was utilized to analyse DNA samples. Gels of 1% [w/v] agarose were prepared by dissolving agarose in Tris acetate EDTA (TAE) buffer (40 mM Tris pH 8.0, 20 mM acetic acid, 1 mM EDTA). SYBR Safe DNA Gel Stain (Invitrogen, CA, USA) was added to the molten agarose, using a ratio of 1:10,000, before the solution was poured into the cast to set. Samples were loaded onto the agarose gel with Blue/Orange 6x Loading Dye (Promega, WI, USA). For comparison either a 1 kb DNA Ladder or the 100 bp DNA Ladder [(New England Biolabs (NEB, MA, USA)] was loaded alongside the DNA samples. These ladders contain molecular weight standards ranging from 0.5-10 kb and 100-1,500 bp, respectively. Electrophoresis was carried out using Bio-Rad apparatus and visualisation was performed under UV light on an InGenius system (Syngene, U.K.).

3.4.2. Cloning

Synthetic genes

cDNA was available for all the constructs discussed in the following chapters.

Plasmid isolation

For plasmid DNA propagation, 10 mL cultures of LB supplemented with the appropriate antibiotic were inoculated from a glycerol stock. The cultures were grown overnight at 37 ^{o}C in a shaking incubator at 200 rpm. Cells were harvested and the plasmid DNA was isolated using the Wizard Plus SV Minipreps DNA Purification System (Promega, WI, USA).

Polymerase chain reaction

Polymerase chain reaction (PCR) consists of the amplification of a specific DNA fragment from a template DNA molecule (Saiki *et al.*, 1988, Saiki *et al.*, 1985).

Modifications of the basic PCR reaction permit the introduction of mutations (deletions, insertions and site-directed mutagenesis) of the gene of interest. The targeted fragment is amplified by a thermostable DNA polymerase, which extends the fragment delimited by two specific oligonucleotides of 20-40 bp, the PCR primers, which anneal to the 5′ and 3′ ends of the fragment.

The DNA polymerase KOD Hot Start polymerase (EMD Millipore, MA, USA) was used. All primers were purchased from Eurofins MWG, (Munich, Germany). PCR reaction contained a combination of PCR primers, template DNA, PCR buffer, dNTP and Magnesium sulfate ($MgSO_4$) was used. Reactions with and without betaine and DMSO were set up. A typical PCR thermal cycle would consist of an initial denaturation phase, denaturation phase (first step of the cycle), annealing phase (temperature determined by the base composition of the PCR primers) and an extension phase (last step of the cycle). A number of cycles between the first and the final step are carried out. The final extension ensures the completion of those fragments not fully polymerised at the end of the last cycle. The specific reaction mixtures and the thermal cycles used for PCR amplifications of the insert are given in the respective chapters. A machine, PTC-100 Thermal cycler (MJ Research, MN, USA) was used for setting up PCR amplification reactions. The samples are analysed by analysing them on an agarose gel as described above. The Wizard SV Gel and PCR Clean-up System (Promega, WI, USA) was used for PCR clean up and to prepare samples for annealing with the cloning or expression vector.

Transformation

The heat-shock method was used to transform ligated DNA into TOP10 competent cells (Invitrogen, CA, USA) according to the manufacturer's instructions. The reactions were plated on LB agar with appropriate antibiotic(s) and incubated overnight at 37 ^{o}C.

Purified and sequence verified (Eurofins MWG Operon, Ebersberg, Germany) DNA plasmids were used to transform expression cell lines according to the manufacturer's instructions, again using the heat-shock method. As before the

reactions were plated on LB agar with appropriate antibiotic(s) and incubated overnight at 37 ^{o}C.

Screening for positive clones

After ligation and transformation into cloning competent cells, such as *E. coli* TOP10 cells, resulting transformants were tested for the presence of the desired insert. Individual colonies of transformants were used to inoculate 10 mL LB cultures [containing appropriate antibiotic(s)] and their plasmid DNA was extracted as described previously. These extracts were then sent for sequencing (Eurofins MWG Operon, Ebersberg, Germany). A test digest was performed on the resulting DNA for non–Ek/LIC clones prior to sending for sequenicng. Each test reaction contained 7 µL DNA, 1 µL FastDigest buffer (10x) and 1 µL each of the FastDigest restriction enzymes. (Invitrogen, CA, USA). The reaction mixture was incubated for one hour and 37 ^{o}C. The digest was analysed by gel electrophoresis.

Glycerol stock preparation

After being transformed with the sequence-verified plasmid DNA, protein expression competent cells were plated on LB agar with appropriate antibiotics and incubated overnight at 37 ^{o}C. A single colony was picked and used to inoculate 10 mL LB with appropriate antibiotic(s) and the culture was grown overnight at 37 ^{o}C with shaking. For each glycerol stock, 500 µL of culture was added to 500 µL of 50% [w/v] glycerol in a cryovial and flash frozen in liquid nitrogen followed by storage at -80 ^{o}C.

3.5. Protein methods

Buffer Preparation

The buffer solutions were prepared by solublisation of the appropriate quantities of the buffer and other required reagents to a small quantity of water. After the reagents were dissolved, the pH of the buffer solution was adjusted using the appropriate acid or base and analysed using the 3510 Jenway pH

meter (Staffordshire, U. K.). Sufficient quatity water was then added to make up to the expected volume. The buffers were filtered using Nalgene Rapid flow filter units (Thermo Fisher Scientific, MA, USA).

3.5.1. Protein expression

Primary cultures (50 mL LB cultures in 250 mL conical flasks) with appropriate antibiotic(s) were inoculated from glycerol stock and grown overnight at 37 ^{o}C in a shaking incubator at 200 rpm. 10 mL of these seed cultures were used to inoculate 1 L secondary cultures, also containing antibiotics, and these were then grown at 37 ^{o}C in a shaking incubator (200 rpm) until optical density (O.D.) reached 0.6-0.9 at 600 nm. Protein expression was then induced by the addition of isopropyl-1-thio-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The incubation temperature was varied if required for expression and induction with IPTG. Cells were grown overnight and harvested by centrifugation. For storage, cell pellets were resuspended in their respective purification buffers and frozen at -80 ^{o}C until the day of use.

Expression trials

Primary cultures (10 mL LB cultures in 50 mL falcon tubes) with appropriate antibiotic(s) were inoculated with glycerol stocks and grown overnight at 37 ^{o}C in a shaking incubator. 1 mL of these seed cultures were used to inoculate 50 mL media in 250 mL conical flasks containing the respective antibiotics. Expression trials in LB and TB media were grown at 37 ^{o}C to an O.D. of 0.6–0.9 and then induced with 1 mM IPTG. Depending on the design of the expression trial, induction was performed at either 37 ^{o}C or 16 ^{o}C. Samples were collected at regular intervals and expression was assessed by SDS-PAGE. Expression trials in AI media were performed by inoculation of the AI media followed by its incubation for 24 h at appropriate temperatures (37 ^{o}C or 16 ^{o}C). All expression trials were carried out in a shaking incubator at 200 rpm. Cells were harvested after induction by centrifugation (4 ^{o}C at 12,000 g, 10 min).

3.5.2. Protein purification

Specific details of the purification process are given in each chapter for each of the proteins purified. A general summary of the experimental procedures used is given below.

Cell preparation

Frozen cell suspensions were thawed and cell lysis was carried out using a cell disruptor (TS Series, Constant System) at 20,000 PSI at 4 ^{o}C.

For soluble protein, the cell lysate was centrifuged at 30,000g for 30 min at 4 ^{o}C to separate the clear fraction. The supernatant was then purified using a range of chromatographic methods.

For insoluble protein, the cell lysate was centrifuged at 10,000g for 30 min at 4 ^{o}C. The resulting inclusion body pellet was then washed in lysis buffer and solubilised before being refolded in respective refolding buffers. The refolded protein was purified using appropriate chromatographic techniques.

Protein Purification

The proteins were purified on the basis of their biochemical properties and the presence of affinity tags. The commonly used purification techniques were immobilized metal ion affinity chromatography (IMAC), ion exchange chromatography (IEX) and size exclusion chromatography (SEC). Carefully designed buffers were filtered prior to use. Details of the chromatographic techniques, columns and media used are given in the following chapters. Purifications were performed as per the manufacturer's instructions for each column and media. All proteins were purified using ÄKTA systems (GE Healthcare, U.K.).

Solvent exchange/ Dialysis

Solvents were fully exchanged by dialysis or gel filtration chromatography using a desalting column between purification stages when required. The

protein was dialysed in Medicell membranes (U.K.) tubing with an appropriate molecular weight cut-off and was dialysed overnight at 4°C against the buffer (100x the sample volume) to be used for the following purification step. The buffers were replaced 2-3 times to ensure full exchange. When large sample volumes were to be solvent exchanged, multiple rounds of desalting using a HiPrep 26/10 desalting column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with the final buffer was used.

Protein solubilisation strategies

For insoluble protein expression, cell lysis and clarification of the cell cultures after expression yielded inclusion bodies. Four solubilisation conditions with guanidine HCl or urea as the chaotrope were used for solubilisation trials [Table 3.2] by stirring the pellet overnight at 4 °C. The degree of solubilisation was analysed by SDS-PAGE and the appropriate condition was then used for large scale experiments.

Table 3. 2: Trial solubilisation conditions.

Buffer	Condition
A	100 mM Tris HCl pH 8, 500 mM NaCl, 7 M guanidine HCl
B	100 mM Tris HCl pH 8, 500 mM NaCl, 8 M urea
C	100 mM Tris HCl pH 8, 500 mM NaCl, 7 M guanidine HCl, 0.15 M glutathione reduced
D	100 mM Tris HCl pH 8, 500 mM NaCl, 8 M urea, 0.15 M glutathione reduced

Protein refolding strategies

The aim of refolding is to favour protein refolding over aggregation. This is done by keeping the protein concentration at or below 0.1 mg/mL. The denatured protein was refolded using the drop wise dialysis method at 4 °C. This method ensures the instantaneous elimination of the chaotrope upon 20-fold (or larger) dilution. The dilution buffer was carefully selected by screening conditions provided in the refolding screen, iFOLD refolding system (EMD

Millipore, MA, USA). The degree of solubilisation was determined by SDS-PAGE.

The solubilised protein was dropped into a the selected refolding buffer (volume 20x larger than the solubilised protein) at a speed of 0.1 mL/min with gentle stirring at 4 ^{o}C. The refolded protein was left undisturbed overnight at 4 ^{o}C and purified the following day using appropriate chromatographic techniques.

Protein concentration determination

The final concentration of the purified protein was determined using a NanoDrop 2000c (Thermo Fisher Scientific, MA, USA) by A_{280} measurements using an extinction coefficient calculated from the protein sequence by ProtParam (Gasteiger *et al.*, 2005). The average of three sample readings was used for the final concentration determination.

3.5.3. Protein identification

3.5.3.1. SDS-PAGE

Protein expression and the subsequent purification steps were analysed by sodium dodecyl sulfhate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions of interest (uninduced samples, induced samples, column load, column wash and elution peaks) were prepared by dilution with SDS-PAGE sample loading buffer 5x (0.25M Tris-HCl, pH 6.8, 15% [w/v] SDS, 50% [w/v] glycerol, 0.01% [w/v] bromophenol blue) with or without 25% v/v β-mercaptoethanol. Samples were denatured by heating for 10 min at 95 ^{o}C before being loaded onto an 8%, 10%, 12% or 15% [w/v] BIS-TRIS gel (selected based on the molecular weight of protein). Depending on sample, 10 to 20 µL was loaded. Protein molecular weight markers (2-212 kDa; NEB, U.K.) were used as standards. Gels were run for 60 min at 160 volts in MOPS running buffer (50 mM MOPS, 50 mM Tris base pH 7.7, 0.1% [w/v] SDS, 1mM EDTA) using a Mini-PROTEAN tetra cell (Bio-Rad, CA, USA).

3.5.3.2. Western blotting

Samples were prepared as per SDS-PAGE analysis with the PageRuler prestained protein ladder (Thermo Scientific, MA, USA) for molecular weight standards. After electrophoresis, gels were set up in a blotting cassette with filter papers and with either a nitrocellulose or a PVDF membrane (Millipore, MA, USA). Blotting was carried out in SDS running buffer (24mM Trizma, 0.19 M glycine, 0.1% [w/v] SDS) with 20% v/v methanol for one hour at 60 volts. The blots were blocked in 5% [w/v] milk powder in 1x Tris Buffer system (TBS) with 0.01% [v/v] Tween (TBST). Stock solution of 10x TBS (200 mM Tris HCl pH 7.6, 1.5 M NaCl) was prepared and diluted as per required. Following a wash step using TBST, primary antibodies, some of which were conjugated with HRP, were diluted appropriately in 5% [w/v] milk powder in TBST and added to the blots. After incubation for one hour at room temperature, with rocking, the blots were washed five times for 10 min each in TBST. HRP conjugated secondary antibodies (if required) were diluted appropriately in 5% [w/v] milk powder in TBST and incubated for one hour with shaking at room temperature. After five washes, blots were revealed using the 3,3'-diaminobenzidine (DAB) colorimetric detection method. The antibodies (Ab) used in this thesis include monoclonal anti-ploy-His Ab conjugated with HRP (referred to as anti-His Ab) , unconjugated anti-GST Ab raised in goat for which anti-goat IgG-peroxidase Ab produced in goat was used as the 2^{o} Ab (referred to as anti-GST Ab), unconjugated anti-SIRT2 Ab produced in rabbit and anti-SIRT1 Ab produced in rabbit for which anti-rabbit IgG-peroxidase Ab produced in goat was used as the 2^{o} Ab (referred to as anti-SIRT2 Ab and anti-SIRT1 Ab, respectively). All antibodies were purchased from Sigma-Aldrich Co., Kent, U.K.

3.5.3.3. Mass spectrometry

Intact protein

The purified protein was dialysed in water at 4 ^{o}C overnight. The protein was then concentrated to between 3 and 10 mg/mL and sent for mass spectrometric analysis. Before analysis, 0.1% v/v acetic acid was added to the sample to facilitate the ionisation of the protein for mass spectrometry. Samples for intact

protein mass spectrometry were performed at the Chemical Characterisation and Analysis Facility, University of Bath.

In-gel tryptic digest mass spectrometry

The purified protein sample was analysed by SDS-PAGE. The protein band of interest was excised and sent to the Proteomics Facility, University of Bristol for in-gel trypsin digest mass spectrometry. The protein samples were trypsin digested using ProGest automated digestion unit which minimises sample handling and keratin contamination. The resulting peptides were then analysed by MALDI-TOF/TOF mass spectrometry to generate peptide mass fingerprints and peptide sequence information, which was then searched against various databases using a Mascot server to identify the protein. Proteins were verified based on a match with the top candidate from the search.

3.6. Biochemical assays

Biochemical assays were used to confirm the identity of the purified proteins and ensure they were biologically active. These included the *Fluor-de-Lys* SIRT2 fluorometric drug discovery assay (Enzo Life Sciences, NY, USA), used to measure the lysyl deacetylase activity of recombinant sirtuins and the isothermal titration calorimetry assay, which was used to measure the binding affinity of galectins to the dendrons used in the study.

These assays have been discussed further in the following chapters.

3.7. Protein crystallography

Protein crystallography has had a profound influence on structural biology. The ability of X-ray crystallography to provide accurate structural details at the atomic resolution has had a huge impact on biological research. The information it provides has become a prerequisite for biological studies such as deciphering mechanisms, understanding activity of proteins, identifying substrates, rational based drug design and many others.

To date, 28 Nobel prizes have been awarded for research towards macromolecular X-ray crystallography. The contribution of X-ray diffraction to biology took off with Rosalind Franklin's and Maurice Wilkin's experiments on the genetic blueprint for life, DNA, which lead to the work of Francis Crick and James Watson in the early 1950's for which they won the Nobel prize. In the same year, the pioneering work by Nobel laureates Max Perutz and John Kendrew on haem-containing proteins, including the structure of myoglobin, was also recognised. Since then X-ray crystallography has contributed significantly to the field of structural biology which is evident in the number of structures deposited in the Protein Data Bank (PDB). Of the 103,354 structures deposited (as of 23rd September, 2014) 91,655 structures have been solved by X-ray crystallography. This technique has now become a gold standard in understanding the function and regulation of biological macromolecules and assemblies. Complemented by the high resolution structural data obtained by Nuclear Magnetic Resonance (NMR) and electron diffraction, this wealth of knowledge provides a fundamental basis for current thinking in structural biology.

The appropriate wavelength of X-rays (in the Ångström range, ~10^{-10} m) makes it suitable for scattering by the electron cloud of an atom. Often, X-rays at a wavelength of 1.54 Å (Cu-Kα, rotating anode home sources) are used for macromolecular crystallography which is very similar to the distance between bonded carbon atoms. Crystals of proteins behave as an amplifier to increase the intensity of diffraction due to their periodicity, *i.e.* repeated unit cells [Figure 3.5]. Thus scattered waves can add up in signal and raise the signal to a

measurable level. As the waves add up in phase in some directions, they cancel out in a lot of other directions. This results in the diffraction pattern of a crystal being an array of spots [Figure 3.3]. The structure of the protein (electron density) is back calculated for many such diffraction patterns collected from the crystal. The typical process is outlined in figure 3.2.

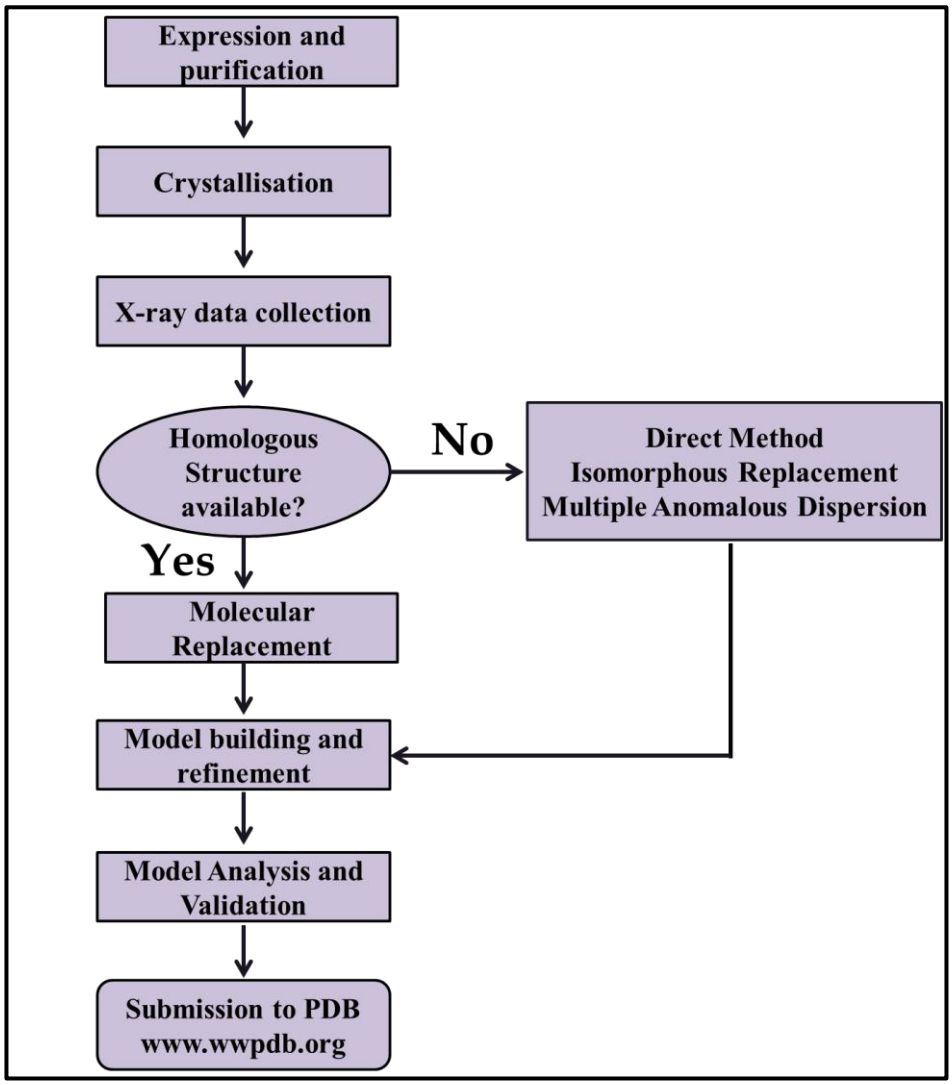


Figure 3. 2: Outline of the steps involved in protein crystallography.

The process of determining macromolecular crystal structures has remained largely unchanged over the years. However, the continued advancement of underlying crystallographic technology and methodology, complemented by the revolution in molecular biology, has driven the significant increase of successes in this field. This development extends to the kind of problems that can be tackled, the time involved and the accuracy and efficiency with which

protein structures can be determined. In spite of this progress, growing crystals that diffract to high resolution still remains one of the major bottlenecks in this field.

As all steps after crystal mounting are now routinely performed by a robot and then on a computer, the advancement in computational resources and performance available has had a huge influence on protein crystallography. The availability of inexpensive computing power has contributed to radical changes in data collection, processing and refinement approaches.

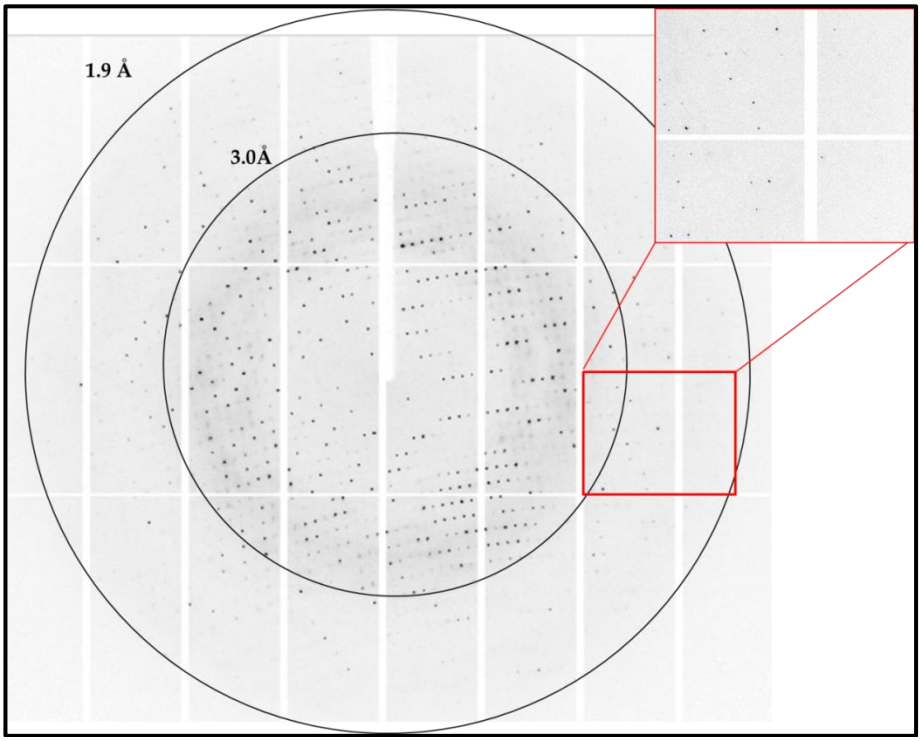


Figure 3. 3: Representative image of a diffraction pattern.

3.7.1. Theory

3.7.1.1. Principle of X-ray crystallography

The wave nature of X-rays was established when W. L. Bragg in 1913, first elucidated the crystal structures of sodium chloride, potassium chloride, potassium bromide and potassium iodide, using X-ray analysis (Bragg, 1913). Crystallography relies on the basic theory of X-ray scattering from a crystal, which is a well understood phenomenon. Crystals can be thought of as a three-dimensional array of identical unit cells. This periodicity leads to Laue equations and Bragg's law.

The three Laue equations give the conditions to be satisfied by an incident wave to be diffracted by a crystal. The Laue diffraction conditions are as follows:

$$\mathrm{a} \cdot \mathbf{S} = h_1$$
$$\mathrm{b} \cdot \mathbf{S} = h_2$$
$$\mathrm{c} \cdot \mathbf{S} = h_3$$

Where,
a, b and c are the crystal translation vectors,
S the vector normal to the reflecting plane with magnitude $2\sin\theta/\lambda$, and
h_1, h_2 and h_3 are integers ……..**(1)**

Bragg's Law (Figure 3.4) provides the condition for a plane wave to be diffracted by a family of lattice planes and result in constructive interference.

$$2\mathrm{d}\,(hkl)\sin\theta = n\,\lambda \qquad \text{……..(2)}$$

Where,
d is the distance between the lattice planes (hkl)
θ is the reflecting angle , n is an integer and
λ is the wavelength [with respect to Laue conditions given in (1)]
$h_1=nh$, $h_2=nk$ and $h_3=nl$; it is common practice to absorb n in h, k and l.

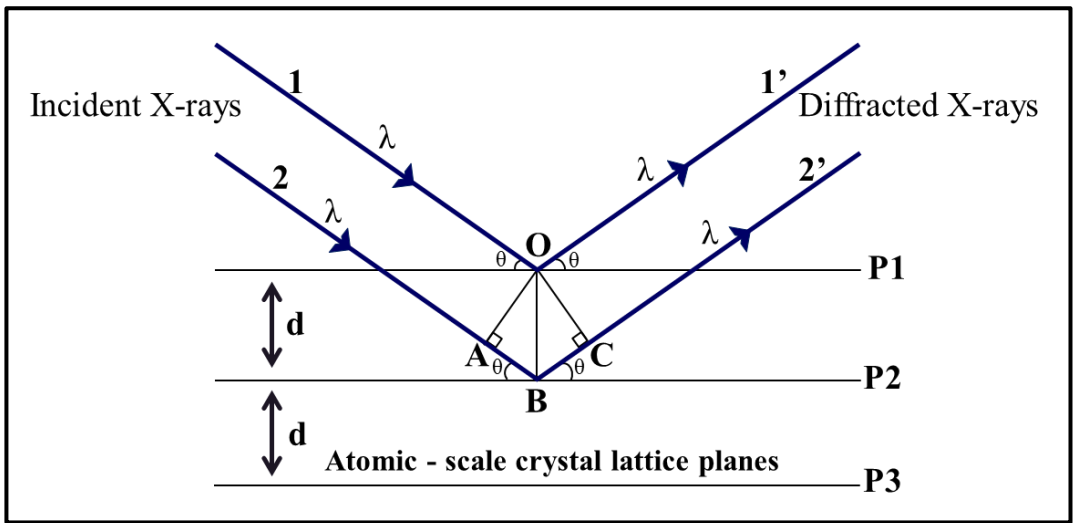


Figure 3. 4: Schematic representation of Bragg's Law.

Two beams (1 and 2) with identical wavelength (λ) and phase approach a crystalline solid which gets scattered off two different atoms (O and B) within it. The lower beam traverses an extra length of 2dsinθ. The diffracted X-rays exhibit constructive interference when the extra distance traversed by beam 2 (ABC) differs by an integral (n) multiple of the wavelength.

A crystal is a regular arrangement of molecules(s), repeated translationally, on a three-dimensional lattice. The asymmetric unit is the smallest portion of a crystal structure to which symmetry operations can be applied in order to generate the complete unit cell [Figure 3.5]. The application of symmetry operations (a combination of rotation and translations for biological macromolecules) to an asymmetric unit yields one *unit cell* which when translated in three dimensions, makes up the entire crystal. The asymmetric unit contains the unique part of the crystal structure which is related to all other identical motifs in the unit cell governed by symmetry. The unit cell can be described by the lengths of three edges (a, b and c) and the angles between them (α, β and γ) [Figure 3.5].

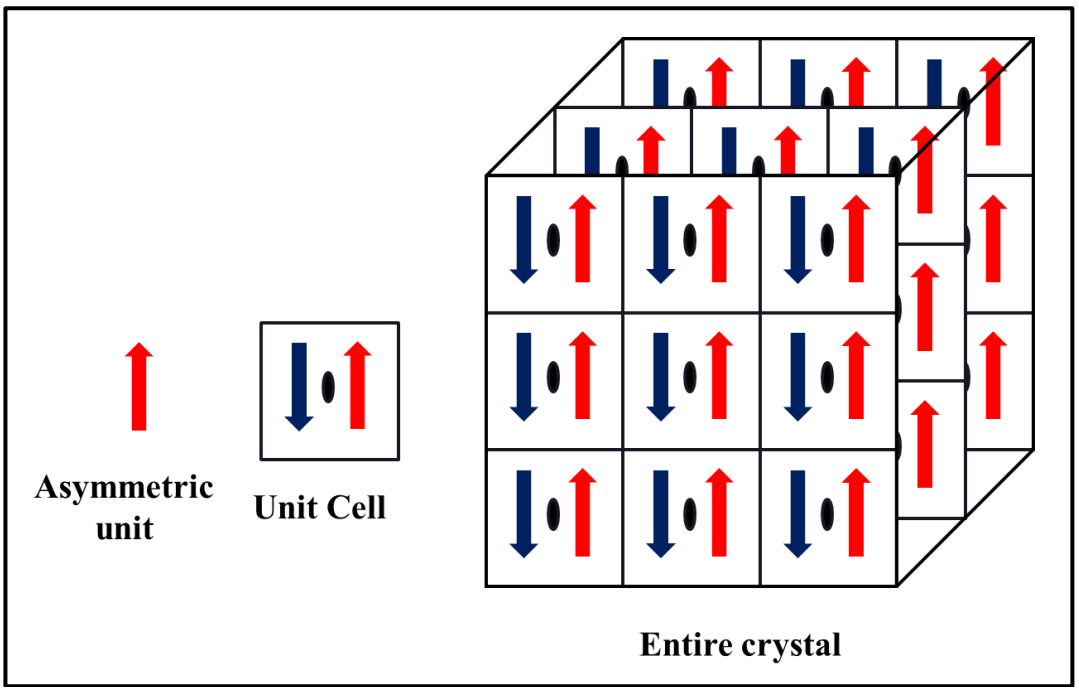


Figure 3. 5: Fundamental parts of crystal structures.

The asymmetric unit (red upward arrow) is rotated 180° about a two-fold crystallographic symmetry axis (black oval) to produce a second copy (blue downwards arrow). The asymmetric unit may contain several protein molecules or just one. Together, the arrows comprise the unit cell. The unit cell is then transnationally repeated in three directions to make a crystal.

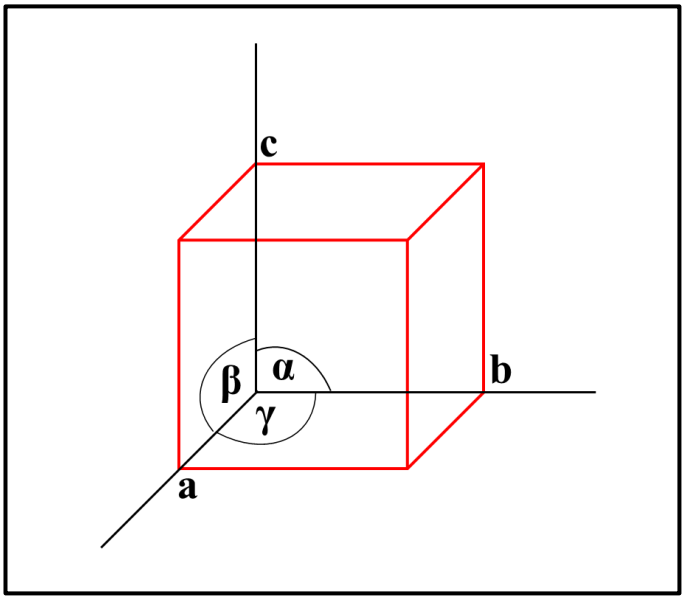


Figure 3. 6: Measurable parameters of a unit cell.

The four basic types of symmetry operations possible for a crystal lattice are rotation, reflection, inversion and translation. Based in the highest symmetry element present, crystals can be grouped into 7 crystal systems and these can be further divided into 14 space lattices called the 'Bravais Lattices' (Blundell and Johnson, 1976), depending on the type of lattice (primitive or non-primitive) chosen to describe the unit cell [Table 3.3]. A combination of these symmetry elements gives rise to 230 space groups, of which only 65 are permitted for proteins.

Table 3. 3: Bravais Lattices

Name	Bravais Lattice types	Lattice Restrictions
Triclinic	P	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$
Monoclinic	P, C	$a \neq b \neq c$; $\alpha = \gamma = 90^o \neq \beta$
Orthorhombic	P, C, I, F	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90^o$
Tetragonal	P, I	$a = b \neq c$; $\alpha = \beta = \gamma = 90^o$
Trigonal	P	$a = b \neq c$; $\alpha = \beta = 90^o, \gamma = 120^o$
	or R	$a = b = c$; $\alpha = \beta = \gamma < 120^o, \neq 90^o$
Hexagonal	P	$a = b \neq c$; $\alpha = \beta = 90^o, \gamma = 120^o$
Cubic	P, I, F	$a = b = c$; $\alpha = \beta = \gamma = 90^o$

Lattice types: Primitive (P), Base-centred(C), Face-centred(F), Body-centred (I).

The resolution of a crystal refers to the minimum interplanar spacing of the reflections included in structure determination. The unit of resolution is the Angstrom (Å = 10^{-10} m). It describes the ability to distinguish between neighbouring features in an electron density map.

3.7.1.2. Structure factors

Crystals only diffract X-ray beams constructively in certain discrete directions. The intensity of the diffracted beam (*hkl*) is proportional to the square of the amplitude of the Structure Factor **F**(*hkl*).

The structure factor is a function of the electron density distribution in the unit cell :

$$\mathbf{F}(hkl) = \mathrm{V} \int_{\text{cell}} \rho\,(x,y,z)\; e^{2\pi i\,(hx+ky+lz)}\, dx\; dy\; dz \qquad \text{.....(3)}$$

Where,
V is the volume of the unit cell and ρ is the electron density distribution.

Electron density distribution in the unit cell can be calculated by an inverse Fourier transform of this function and by virtue of its periodicity expressed in a summation as :

$$\rho\,(x,y,z) = \frac{1}{\mathrm{V}} \sum_{h} \sum_{k} \sum_{l} \mathbf{F}(hkl)\; \mathrm{e}^{-2\pi i\,(hx+ky+lz)} \qquad \text{.....(4)}$$

The structure factor in equation (3) is a complex quantity:

$$\mathbf{F}(hkl) = \left|\mathbf{F}(hkl)\right| \mathrm{e}^{i\varphi\,(hkl)} \qquad \text{.....(5)}$$

The amplitudes of $|\mathbf{F}(hkl)|$ can be derived from the experimentally measured intensities of the diffracted beam (*hkl*). However, the phase angles φ(*hkl*) cannot be obtained directly from the diffraction pattern. Without this information, equation (4) cannot be solved. This is known as the 'phase problem' in crystallography. Several methods can be used to determine the phases indirectly (Rupp, 2010). These include the direct and the indirect methods.

Direct methods

Direct methods were developed by Hebert Hauptman and Jerome Karle for which they were awarded the Nobel prize in 1985 (Karle and Hauptman, 1956). In this method, the relationship among phases, based on prior knowledge about

crystal structures like atomicity and non-negativity of ρ(xyz) can be used to evaluate the phases from the measured intensities. If it is assumed that the crystal is made-up of similarly shaped atoms and that all atoms have positive electron density, then statistical relationships between sets of structure factors can be established. These relationships can be used to deduce possible values for phases. Direct methods exploit such relationships and can be used to solve the structures of small molecules. However, the statistical relationships become weaker as the number of atoms increases, and direct methods are limited to structures, with at most a few hundred atoms in the unit cell. Although there are developments that push these limits, particularly for crystals that diffract to very high resolution (1.2 Å or better), direct methods are not used frequently in macromolecular crystallography.

3.7.1.3. Indirect methods:

Molecular replacement (MR)

When a homology model is available, molecular replacement can be successful using methods first described by Michael Rossmann and David Blow (Rossmann and Blow, 1962). In MR, a model structure is generated from a crystal structure or solution structure for a homolog of the unknown structure. The method works mainly by computing the three rotation angles to correctly orient the model and then by searching for the translation vector, to place the search model in the unit cell of the unknown protein. Once the search model is correctly oriented, it needs to be placed within the unit cell relative to the symmetry of the crystal. Doing a rigid body refinement can optimise the solution so obtained and *R-factor* or the *Reliability index* can be used to gauge the validity of the structure solution. The R-factor is an estimate of error between the observed and the calculated structure factors. The correct solution will give the lowest R-factor. With the increasing knowledge of three dimensional structures, the accuracy, reliability and success of this method has also increased.

Multiple isomorphous replacement (MIR)

Max Perutz and his co-workers pioneered the use of isomorphous replacement (Green *et al.*, 1954). This method was the first phasing technique which became available to macromolecular crystallographers and is widely used for initial phasing when a close structural homologue of the protein in question is unavailable. This method involves the isomorphous attachment of heavy atoms to protein molecules in the crystal, generally by soaking the crystal in a solution containing the heavy atom of choice. In this way, a few heavy atoms replace some water molecules without changing the unit cell dimensions or the structure of the protein itself. On comparing the data collected and intensity differences from the native crystal and the heavy atom labelled crystal, positions of the heavy atoms and in turn the phase angels can be calculated to overcome the phase problem. However, isomorphous replacement has several problems including non-isomorphism between crystals (changes in unit-cell, reorientation of the protein, conformational changes, changes in salt and solvent ions), locating of heavy atoms, problems in refining heavy atom positions, occupancies and thermal parameters as well as errors in intensity measurement.

Multiple wavelength anomalous dispersion (MAD)

Like isomorphous replacement, this technique also exploits the abrupt changes to the scattering power of heavy atoms near their absorption edges (Hendrickson *et al.*, 1985). MAD overcomes the problem of non-isomorphism observed in isomorphous replacement. Data are collected at several wavelengths, typically three, in order to maximise the absorption and dispersive effects. At certain wavelengths, tightly bound electrons in an atom do not scatter as free electrons do and thus cause an anomalous scattering effect. This effect depends on the wavelength and is generally stronger for heavy atoms. This anomaly results in unequal intensities of otherwise equivalent reflections, called Bijvoet pairs. This can be used to calculate the phase angle and facilitates phase determination from a single crystal.

3.7.1.4. Phase improvement

The degree of success of the phasing experiment is affected by the quality of the diffraction data and the model for molecular replacement, or heavy atom model for isomorphous replacement or anomalous dispersion. This in turn can result in large errors in the phases and electron density maps. Several methods have been developed to improve the phases. These methods result from our prior knowledge of a good electron density map and include methods such as density modification by solvent flattening, averaging and histogram matching or by automated fitting and refinement.

In a protein crystal, roughly half the volume is occupied by well-ordered protein molecules and the other half by disordered solvent which have flat featureless electron density. However, this disordered solvent region may contribute to features which will result in phase errors. Solvent flattening will lead to modification of the density making the corresponding phases more accurate. The programmes written for solvent flattening, have algorithms written for automatically defining boundaries between the ordered protein and disordered solvent regions.

As more than one copy of the protein molecule can crystallise in one asymmetric unit of the unit cell and because the same protein appears in different parts of the electron density map, averaging the density will result in elimination of random errors and increase the accuracy of the corresponding phases.

Histogram matching takes advantage of the fact that proteins are made up of similar atoms and relative distances which result in similarity in the density values in the electron density maps of different proteins. Phase errors may be observed due to incorrect distribution of high and low densities which can be altered to the expected distribution based on the prior knowledge of density values of similar atoms in different proteins using a programme for histogram matching. This method would also contribute to increasing the accuracy of the phases.

Similarly, automated fitting and refinement can be done by utilising prior knowledge of bond lengths, bond angles and chemical connectivity defined by the amino acid sequence to build an atomic modelling and iterative refinement.

With the phase problem solved and improved, the initial model can be built in the electron density maps. The model can then be refined against observed X-ray intensity data and standard protein geometrical information. The quality of the final model is assessed through cross validation of the X-ray data and extensive validation of the geometry of the protein model. The resolution of the diffracted data collected, which is determined by the quality of the crystal, plays a key role in the accuracy of the structure determination process.

3.7.2. Crystallisation

Proteins, like many other molecules, can form crystals when the solution in which they are dissolved becomes supersaturated. Under such conditions, protein molecules get packed in a repeating array with non-covalent interactions holding them in place (Russo Krauss *et al.*, 2013). The techniques used for setting up crystallisation experiments of the proteins discussed in this thesis include vapour diffusion (hanging and sitting drop) and microbatch under oil.

The crystallisation of a protein requires a specific environment which is empirically determined for each protein. During crystallisation, the soluble state is slowly transformed towards precipitation and eventually when all conditions are met, the protein molecules arrange themselves in regular orientations to form repeating unit cells. Figure 3.7 illustrates the conditions required for nucleation and crystal growth.

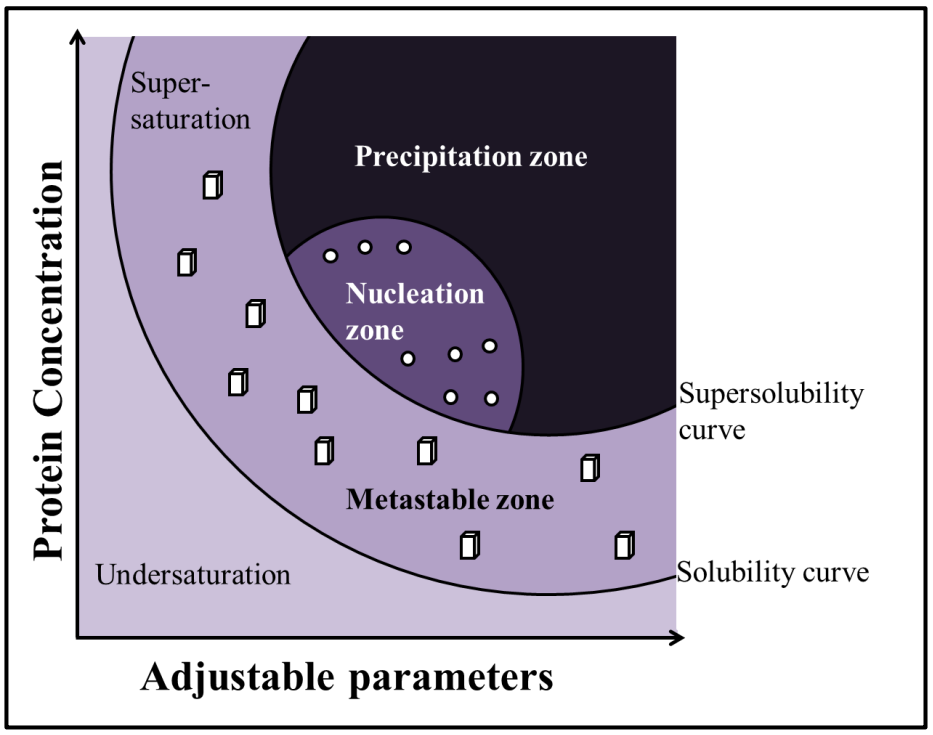


Figure 3. 7: Schematic representation of a protein crystallisation phase diagram.

The horizontal axis can be any crystallising parameter being varied (usually precipitant concentration). The supersolubility curve is the line separating conditions under which spontaneous nucleation (or phase separation or precipitation) occurs from the condition under which the crystallisation solution remains clear if left undisturbed. Theoretically, the more supersaturated the protein solution in the drop, the greater the likelihood that a critical nucleus will form and the smaller the nucleus needed to induce crystal formation. This is represented in the phase diagram by dividing the supersaturated zone into regions of increasing probability of nucleation and precipitation.

Vapour diffusion

Vapour diffusion is the most common method employed for protein crystallisation. In this method, a droplet containing purified protein and the reservoir solution, which consists of salt(s), buffer(s) and/or precipitant(s), are allowed to equilibrate with a large reservoir containing similar salt(s), buffer(s) and/or precipitant(s) in higher concentrations. This allows the water to vaporise from the drop into the reservoir, thus gradually increasing the protein and precipitant concentration in the drop. If the appropriate crystallisation solutions are used for a given protein, crystal growth will occur in the drop on

reaching equilibrium. Vapour diffusion can be performed in either a hanging-drop, a sitting-drop format or by modified microbatch screening under oil [Figure 3.8].

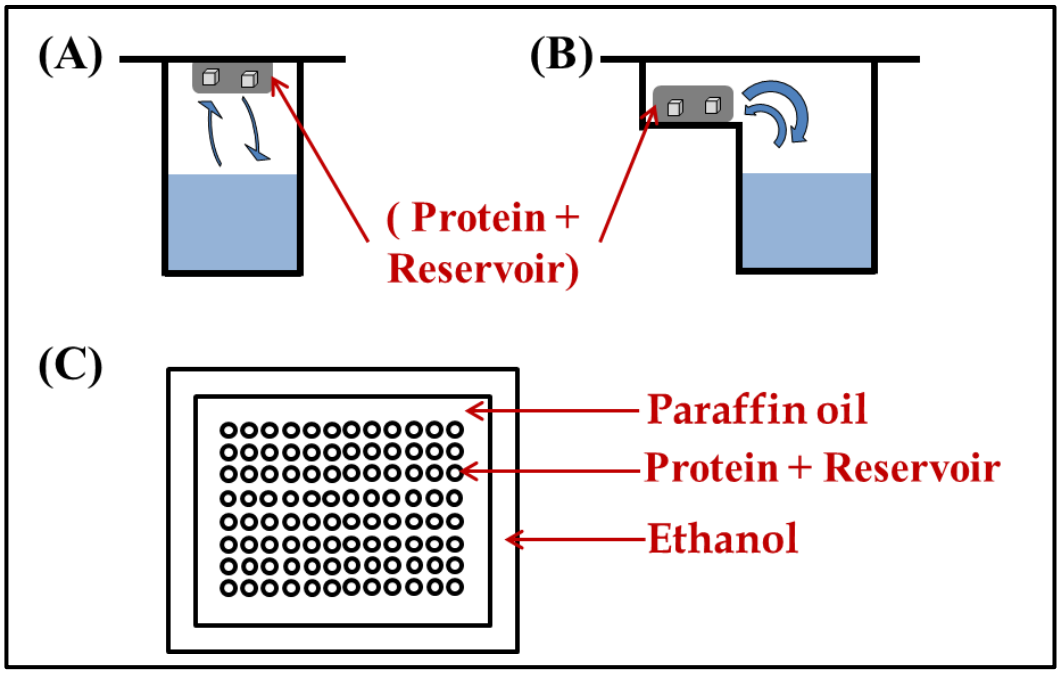


Figure 3. 8: Vapour diffusion method.

(A) Hanging-drop method: A drop of protein and reservoir solution is placed on an inverted cover slip, which is then suspended above the reservoir. (B) Sitting-drop method: A drop of protein and reservoir solution is placed on a pedestal that is separated from the reservoir. Both methods require sealing of the environment for equilibration to occur between the drop and the reservoir. (C) Modified microbatch crystallisation or crystallisation under oil: A drop of protein and reservoir solution is pipetted into the drop and layered with oil. Alcohol is added to the periphery. Oil acts as barrier between alcohol and the drop resulting in gradual diffusion of alcohol into the drop. Vapour batch plates were designed and provided by Douglas Instruments, U.K.

Pre-crystallisation test

A pre-crystallisation test (PCT) was carried out to determine the optimum protein concentration for crystallisation. Table 3.3 gives the details of crystallisation conditions prepared, as recommended in the PCT User Guide (Hampton Research, CA, USA). The hanging drop vapour diffusion method was used to set up a 4-well crystallisation plate. A drop size of 2 µL was used, in which the ratio of protein to reservoir was 1:1. Various protein concentrations

were tested. The PCT plates were incubated overnight at 16 ^{o}C before the degree of precipitation was analysed using a light microscope. The PCT User Guide was followed to determine the initial protein concentration for setting up automated screening.

Table 3.3: Pre-crystallisation test reagents

Reagent	Condition
1	0.1 M Tris-HCl pH 8.5, 2.0 M Ammonium sulfate
2	0.1 M Tris-HCl pH 8.5, 1.0 M Ammonium sulfate
3	0.1 M Tris-HCl pH 8.5, 0.2 M Magnesium chloride hexahydrate, 30% [w/v] PEG 4000
4	0.1 M Tris-HCl pH 8.5, 0.2 M Magnesium chloride hexahydrate, 15% [w/v] PEG 4000

Automated preliminary crystallisation screen : Pheonix

Automated crystallisation screening was used to identify suitable crystallisation conditions for all purified protein. An automated nano-dispenser (Crystal Phoenix, Art Robbins Instruments) was used to set up plates, quickly and accurately. 96-well INTELLI-plates (Art Robbins Instruments) were set up following a sitting-drop vapour diffusion method.

For this set up, 40 µL mother liquor in the reservoir and 200 nl drop size with 2:1, 1:1 and 0:1 (blank drop) protein to mother liquor ratios were used. Several commercially available screens from Molecular Dimensions, U.K. were used: Structure Screen I+II HT-96; Clear Strategy I HT-96; Clear Strategy II HT-96; PACT *premier* HT-96; JCSG-*plus* HT-96; Morpheus HT-96; PGA Screen HT-96; Heavy + Light twin pack HT-96; ProPlex HT-96; The Stura Footprint Combination HT-96; MIDAS HT-96; MemGold HT-96.

Automated preliminary crystallisation screen : Oryx4

An automated nano-dispenser was employed to set up modified microbatch plates (Oryx 4, Douglas Instruments, U.K.). 96-well vapour batch plates

(Douglas Instruments, U.K.) were set up following a modified microbatch vapour diffusion method. Commercial screens, PACT *premier* HT-96, JCSG-*plus* HT-96, Morpheus HT-96, PGA Screen HT-96 and Heavy + Light twin pack HT-96 screens were used for this purpose.

All plates (set up using Oryx and Pheonix) were incubated at 16 ^{o}C. Crystallisation plates were checked immediately after set up followed by regular monitoring after 24 h, 1 week, 2 weeks and then regularly until the drops dried. When nucleation or potential crystallisation hits were observed further optimisation trials were performed. This involved either manually setting up 24-well plates or using automated set ups on the robot.

Manual crystallisation optimisation trials

Any signs of nucleation or crystallisation were used as a starting point for optimisation trials. Trials by hanging drop method using 24-well plates (XRL plates, Molecular dimensions) were performed. Various parameters were used for optimisation trials which included a broad range of precipitant concentrations (0–200% of the initial concentration), buffer systems for pH optimisation and protein to mother liquor ratios (1:1, 1:2, 2:1, 3:1), drop size (1 – 4 μL) and reservoir volume (400-1000 μL).

Seeding

Tiny crystals or part of a crystal or crystalline material, known as the crystal seed, can act as a template for the assembly of macromolecules. Seeds can be used to initiate nucleation and also reduce the time for random molecular collision or interaction which results in nucleation. Modifying the seed quantity and selecting a suitable method can help yield larger crystals with better morphology and diffracting capability. Seeding methods such as streak seeding, microseeding and random microseed matrix seeding were used to optimise crystals (D'Arcy *et al.*, 2007). The methods are described below.

***Streak seeding*:** A probe, usually a cat's or rabbit's whisker, is used to touch an existing crystal to dislodge small fragments, known as crystal seeds, from it. These seeds are then transferred into a pre-equilibrated drop containing the same conditions the seed were taken from, by rapidly running the whisker in a line through the drop. This method should provide seed nucleated crystal growth along the line in the drop.

Microseeding: In this method, pre-existing crystals are crushed by vortexing the crystals with a small glass bead in an Eppendorf to create a seed stock. This stock is then diluted in series and re-introduced back into a pre-equilibrated droplet. The crushed crystals aid nucleation and crystal growth.

Random microseed matrix screening: The seed stock is prepared in a similar way to microseeding. However, this seed stock is then introduced into random commercial screens using the robot. Many salt crystals grow by this method but a few cases of protein crystal growth have been reported (D'Arcy *et al.*, 2007).

Macroseeding: A suitable crystal of good morphology is removed from a well and is washed repeatedly in stabilizing solution. The crystal is then transferred to a pre-equilibrated drop where in this case the crystal is expected to grow in size rather than stimulate in new crystal growth.

Epitaxial nucleation: Impurities like cellulose fibres sometimes provide surfaces that are suitable for macromolecular crystal growth.

Details for the crystallisation screening and optimisation set up are given in each chapter.

3.7.3. X-ray diffraction data collection and processing

X-ray sources

X-ray sources are of two types: Laboratory sources and synchrotron sources. The X-rays are produced in laboratories either by sealed tube sources or by copper-rotating anode (fixed wavelength of 1.5418 Å for Cu-Kα radiation).

Although they are convenient and less expensive, they produce about 100-1000 times less flux than a synchrotron. A synchrotron is a source of electromagnetic radiation usually produced using bending magnets, undulators and/or wigglers. The advantage of a synchrotron source is that it generates tuneable and intense X-rays. Special beam-lines for macromolecular crystallography are available in synchrotrons such as Diamond Light Source (DLS), Oxfordshire, U.K. All diffraction data discussed in this thesis, were collected at DLS macromolecular crystallography beam-lines I02, I03, I04, I24 or I04-1.

Data collection

Crystals were mounted on cryoloops (Molecular Dimensions, U.K.) and flash frozen in liquid nitrogen prior to data collection. The crystal loops were mounted on the goniometer by a robot, centred and tested for X-ray diffraction. When diffraction spots were observed, appropriate settings were applied for data collection (wavelength, exposure time, resolution, oscillation range, beam transmission, detector distance) and images were collected. Depending on the unit cell dimensions and radiation damage of the crystals, images were collected. Beamlines I02, I03, I04 and I04-1 are presently equipped with PILATUS 6M, 6M-F, 6M-F and 2M detectors (Dectris), respectively. The microfocus, beam-line I24 was equipped with PILATUS 3 6M detector (Dectris).

Data processing

X-ray diffraction data processing proceeds through indexing, pre-refinement of crystal orientation, intensity integration, post refinement and scaling. Data were processed using auto-indexing and data reduction programs, MOSFLM (Battye *et al.*, 2011, Leslie, 2006) or XDS (Kabsch, 2010) followed by Xia2 (Winter, 2010) and scaled using AIMLESS (Evans and Murshudov, 2013). Details on data collection and processing are given in the following chapters.

3.7.4. Structure determination, refinement and validation

Molecular replacement

Molecular replacement was used to determine the structures discussed in this thesis. The details are given in the respective sections. Initial phases were obtained using PHASER (McCoy *et al.*, 2007) and the co-ordinates were obtained from the appropriate native PDB files.

Refinement and validation

Crystallographic refinement is carried out to improve parameters of the model structure such that the structure factors calculated from the built model gets closer to the observed amplitudes. Iterative cycles of refinement optimise the atomic model, allow for improved phases and help compute electron density maps that show finer details of the molecular structure.

Refinement of the crystal structures discussed in this thesis were carried out with PHENIX (Adams *et al.*, 2010) and manually with COOT (Emsley *et al.*, 2010) version 0.7. PHENIX, a refinement program, can perform rigid body, and restrained or unrestrained refinement against X-ray data. It minimises the coordinate parameters to satisfy either a maximum likelihood or least square residual. After initial rigid body refinement on the molecular replacement solution, several rounds of restrained refinement were carried out. Between refinement rounds, manual adjustments and model fitting were performed with COOT using real space refinement and regularisation options. COOT was used to add ligand molecules where appropriate and PHENIX was used to automatically refine water molecules. The water molecules were checked using COOT.

Quality of the final structures was assessed by the figure of merit, R factors, plotting of Ramachandran plots and geometric statistics. All structures were checked using MOLPROBITY (Chen *et al.*, 2010) for validation of the structure. Bond lengths and bond angels of the main chain and side chains, planarity, chirality, solvent content and torsion angles of the main chain (φ, ψ), were

monitored by this programme. Any distorted geometry was corrected manually using COOT followed by a round refinement and validation using MOLPROBITY.

The primary measure of refining structure is given by the residual factor or reliability factor or R_{work}. It is a measure of the agreement between the crystallographic model and experimental diffraction data. It is defined by the following equation:

$$R_{work} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|} \qquad \text{.....(6)}$$

Where,
F_{obs} is derived from the measured reflection intensity and
F_{calc} is the corresponding structure factor obtained from the current model.

Another parameter which states the quality of a structure and helps avoid over-fitting of the model is R_{free} (Brunger, 1992). 5-10% of the total reflections are set apart during scaling. Refinement is carried out against the remaining data. R_{free} is calculated against this subset of reflections that is not used for refinement. This helps to reduce model bias and over fitting of data. Iterative model fitting and refinement cycles results in model improvement and both R_{work} and R_{free} values dropping. Decreasing R-factor values are complimented by another parameter called *Figure of Merit* (FOM) which is indicative of how well the model represents true structure factors.

Another vital parameter that helps assess the quality of the model is the Ramachandran plot which underline the stereochemistry, conformation (planarity of the peptide bond, bond angels, bond length, etc.) and the backbone of the conformational angles (φ, ψ) of the current model based on the expected values for simple organic molecules (Ramachandran *et al.*, 1963).

3.7.5. Structure analysis and PDB deposition

On completion of structure refinement and validation, the structure co-ordinates are deposited with the Protein Data Bank (PDB) (Bernstein *et al.*, 1977). The PDB is a vast global repository of all three-dimensional structures elucidates so far.

Analysis of the three-dimensional structure of the proteins and its complexes provides understanding of its functional aspects. The final refined structure reveals the architecture of the protein and its salient features such as its evolution, mechanism of action, basis of interactions among others. Several software packages including PYMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC) for drawing figures and analysis, SUPERPOSE (Krissinel and Henrick, 2004) for RMS calculations and LIGPLOT+ (Laskowski and Swindells, 2011) for ligand binding interactions can be used.

Chapter 4. MULTIVALENT GALACTOSE BASED DENDRIMER RECOGNITION BY HUMAN GALECTIN-7

4.1. Introduction

Human Galectin–7 (hGal-7), is expressed with the first onset of epidermal stratification (Magnaldo *et al.*, 1998) and thus marks the differentiation levels of keratinocyte. It is involved in the migration of epithelial cells and re-epithelisation of corneal and/or epidermal wounds (Cao *et al.*, 2003). Presence of hGal-7 in the areas of cell-cell contact, especially in the upper layers of the epidermis and their down regulation in transformed keratinocytes, suggests its therapeutic value due to its involvement in physiological cell growth through modulation of cell-cell and cell-matrix interactions. Due to its bridging function, the formation of galectin oligomers through supramolecular assembly is an essential element of galectin function, which relies on the multivalence of galectins or ligands (Nagae and Yamaguchi, 2014).

Galectin-7 is a prototype galectin and is known to dimerise in solution. In 1998, Leonidas *et al.* (1998) reported the first crystal structure of hGal-7 in which dimerisation was observed. Morris *et al.* (2004) later confirmed the relevance of the dimer complex using ultracentrifugation techniques and the dimer was also observed in the NMR structure reported by Nesmelova *et al.*, 2012 (Nesmelova *et al.*, 2012). Recent analysis by NMR, circular dichroism and molecular dynamic simulations showed that ligand binding to galectin-7 caused long-range effects resulting in the stabilization of the dimer and exposed positive cooperativity (Ermakova *et al.*, 2013).

To date, the crystal structures of hGal-7 in complex with galactose (PDB: 2GAL), galactosamine (PDB: 3GAL), lactose (PDB: 4GAL), N-acetyl lactosamine (PDB: 5GAL) (Leonidas *et al.*, 1998) and with a galactose-benzyl phosphate inhibitor (PDB: 3ZXE) (Masuyer *et al.*, 2012) have been reported which provide a comprehensive understanding of the carbohydrate recognition mechanism. Furthermore, with its inherent self-assembling capability, hGal-7 represents an ideal prototype lectin to study the formation of ligand-associated supramolecular assemblies.

Dendrimers are synthetic, monodisperse macromolecules consisting of three distinguishing structural features: the core, the branches or arms of the molecule and the termini which can all harbour a variety of chemical functions (Newkome *et al.*, 2001, Villalonga-Barber *et al.*, 2008, Vögtle *et al.*, 2009). The polyvalent nature of dendrimers and their flexibility make them unique nanodevices for biochemical applications (Boas *et al.*, 2006, Gardikis *et al.*, 2012).

In this study, synthetic galactose-based dendrimeric units, or dendrons, of varying size and composition are empolyed to investigate the cross-linking of hGal-7 through carbohydrate recognition to form supramolecular assemblies. The routes for the preparation of these compounds were based on a modular approach, making use of "click chemistry" protocols, whereby various cores, branching units and terminal groups could be combined in different ways to provide a library of dendrimeric structures. The six compounds used in this study therefore constitute a part of this library (Figure 4.1 and 4.2).

This chapter gives details about the expression, purification, crystallisation and isothermal titration calorimetry studies of hGal-7. Diverse crystallisation screens were set up for the dendrons and further optimisation of these conditions was required. These optimisation set ups are discussed in (Section 4.3.2). Crystallisation with all six ligands was attempted. However, X-ray data were collected for only three complexes; *viz*. D1, D2, D3 (Figure 4.1). This data has been analysed (Section 4.3.3) and the results are discussed (Section 4.4). Isothermal titration calorimetry studies were performed to understand the binding affinity of the multivalent dendrons to hGal-7. Various controls, followed by D1 and D2 were used as the first set of ligands to be tested by this method. The preliminary results obtained from these experiments are presented (Section 4.3.4).

The aim of this chapter is to understand the ligand-associated supramolecular assembly formation of hGal-7 using galactose based multivalent dendrimers.

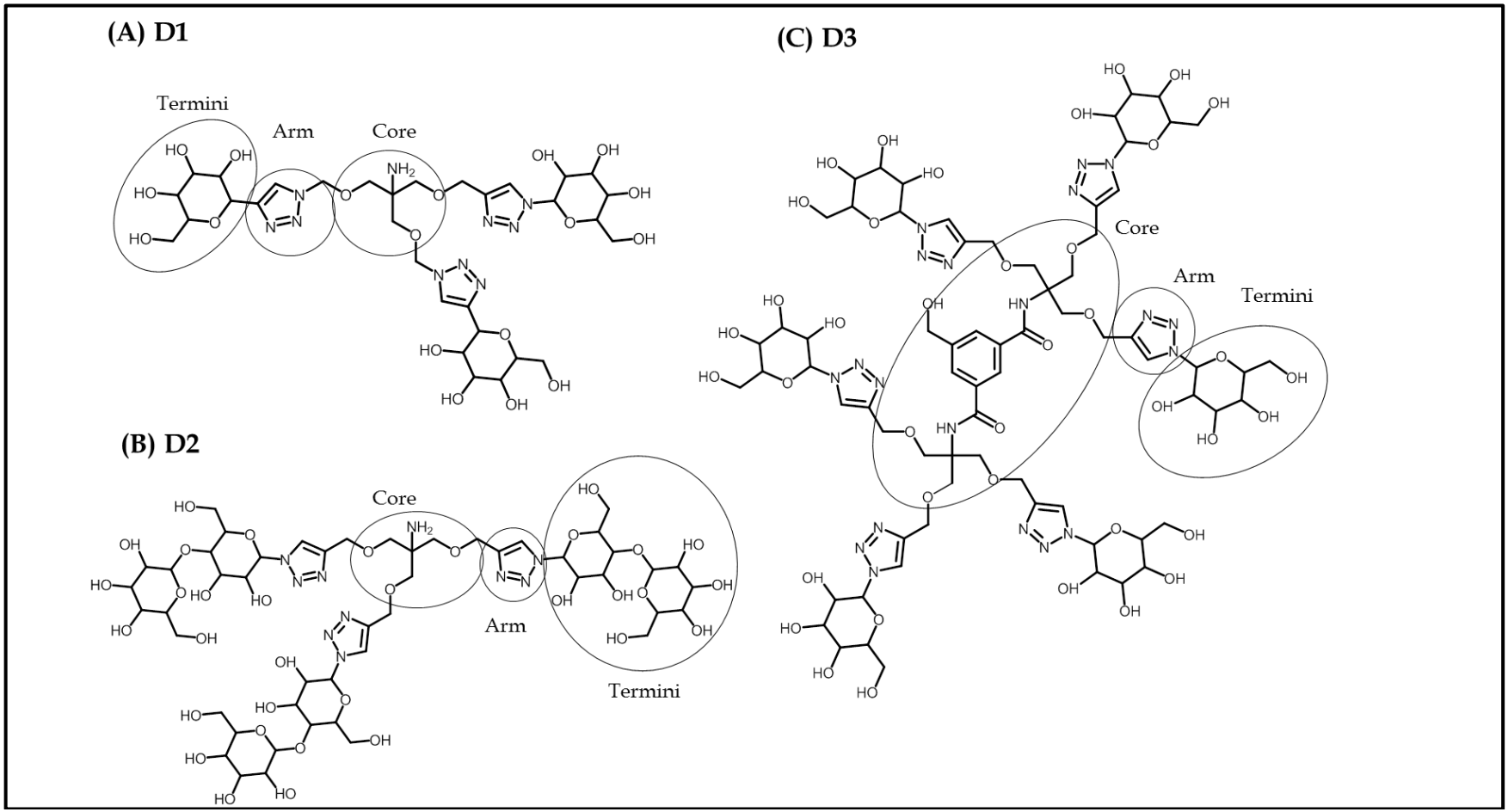


Figure 4. 1 Structures of dendrimers D1, D2 and D3.

The chemical structures of dendrimers with the distinguishing structural features (*i.e.* core, arms and termini) indicated. (A) D1 (B) D2 (C) D3

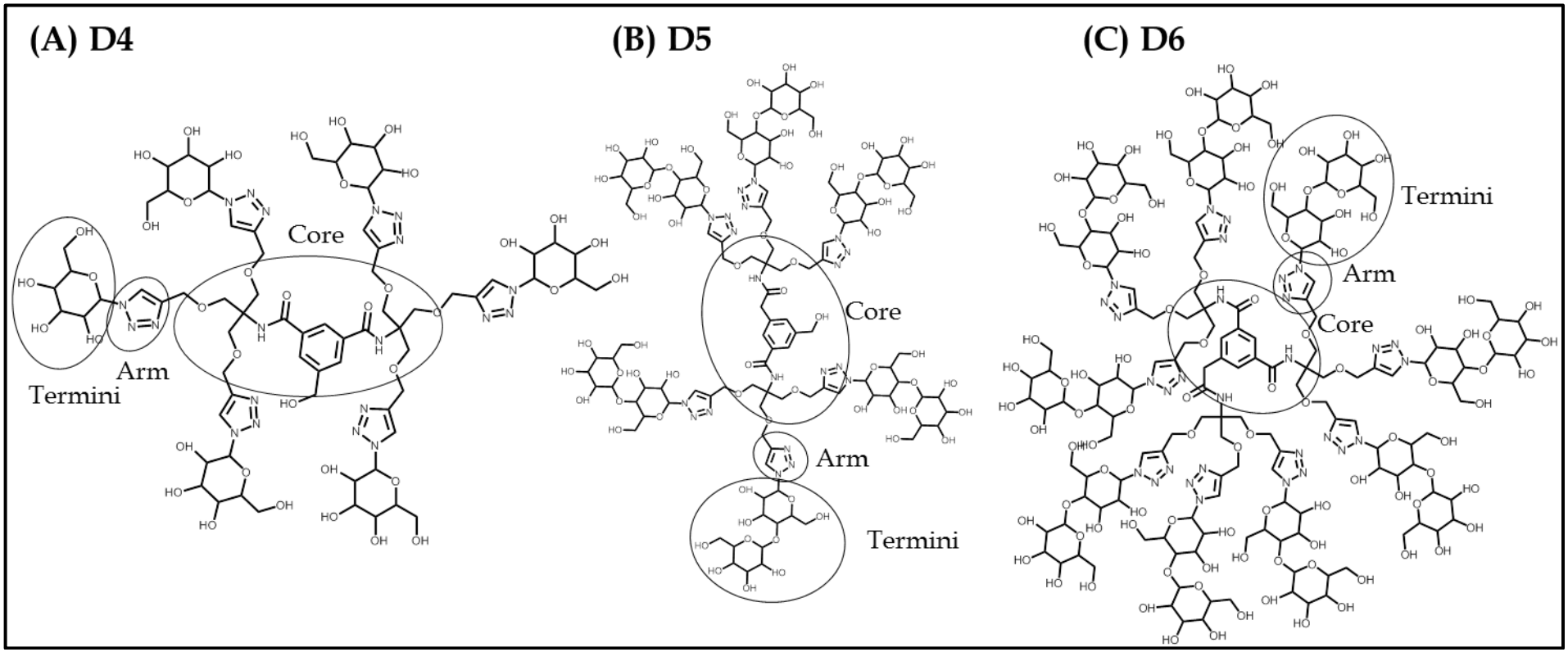


Figure 4. 2: Structures of dendrimers D4, D5 and D6.

The chemical structures of dendrimers with the distinguishing structural features (*i.e.* core, arms and termini) indicated. (A) D4 (B) D5 (C) D6

4.2. Methods

4.2.1. Cloning, expression and purification of hGal-7

The gene coding for hGal-7 was obtained from RZPD (#IRAKp961G19141Q2, Berlin, Germany) and cloned into a modified pET-22 vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal poly-histidine tag. This recombinant plasmid was already available in the laboratory. The plasmid was transformed into *E. coli* BL21 (DE3) cells. Cells were grown in LB media containing ampicillin at 37 °C to an O.D. of 0.6–0.7 and were then induced with 1 mM IPTG at 18 °C. Cells were harvested after 18 h of induction by centrifugation (4 °C at 12,000 *g*, 10 min). Induction of protein expression was analysed by SDS-PAGE.

Cell pellets were resuspended in wash buffer (50 mM sodium phosphate pH 8.0, 200 mM NaCl, 10 mM imidazole) and lysed in an One shot model. The preparation was then clarified by centrifugation (4 °C at 75,000 g, 25 min). The supernatant was loaded on a 5 mL HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer. The column was washed extensively and a single step elution (50 mM sodium phosphate pH 8.0, 200 mM NaCl, 300 mM imidazole) was carried out.

The eluted fraction was dialysed overnight in wash buffer at 4 °C and loaded on a gel filtration column (Superdex 200; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM sodium phosphate pH 8.0, 200 mM NaCl).

4.2.2. Synthesis of the dendrimers

Ligands were synthesised and provided by Dr. Barry Steele's group, National Hellenic Research Foundation, Institute of Biology, Medicinal Chemistry and Biotechnology, Athens, Greece.

The detailed atom labelling for the ligands are shown in appendix C.

Table 4. 1: Chemical formulae and molecular weights of the dendrons.

Dendrons	Chemical Formula	Molecular weight
D1	$C_{31}H_{49}N_{10}O_{18}$	849.78
D2	$C_{49}H_{80}N_{10}O_{33}$	1337.21
D3	$C_{71}H_{102}N_{20}O_{39}$	1861.7
D4	$C_{10}H_{150}N_{30}O_{57}$	2708.45
D5	$C_{107}H_{164}N_{20}O_{69}$	2834.54
D6	$C_{156}H_{240}N_{30}O_{102}$	4167.71

4.2.3. Crystallisation

Pre-crystallisation tests were performed to determine the initial protein concentration for each batch of hGal-7 purified. Automated crystallisation screens of hGal-7 in complex with dendrons were then set up as described in Section 3.6.2. For this purpose, hGal-7 (at a concentration of 10 mg/mL) was pre-incubated at room temperature with the dendron, at protein to dendron ratios of 1:1, 1:2 and 1:3. Different conditions gave crystallisation hits for hGal-7 in complex with the dendremeric molecules. As discussed in Section 3.6.2, the crystallisation screens normally vary in buffer type and pH, salt type and concentration, as well as precipitant type and concentration. The conditions were then optimised by setting up sparse matrix screens around the conditions in which hits were observed. Temperature, protein concentration and various crystallisation techniques are also known to effect crystal growth. The manual optimisation techniques have been previously discussed in Section 3.6.2 and further specific details have been discussed in the results in Section 4.3.2.

4.2.4. Data collection and structure determination

Data were collected at the Diamond Light Source (Oxford, U.K.) on beam-lines I04, I04-1 and I03 which were equipped with a Pilatus 6M detector, Pilatus 1.2M detector and Pilatus 6M detector respectively.

The data were processed using either MOSFLM (Battye *et al.*, 2011, Leslie, 2006) or XDS (Kabsch, 2010) followed by Xia2 (Winter, 2010) and scaled using AIMLESS (Evans and Murshudov, 2013). Initial phases were obtained by the molecular replacement method using PHASER (McCoy *et al.*, 2007) with the co-ordinates of hGal-7 (PDB: 1BKZ) (Leonidas *et al.*, 1998). Crystallographic refinement was carried out using PHENIX (Adams *et al.*, 2010). Model fitting was done using COOT (Emsley *et al.*, 2010) version 0.7. The programme MOLPROBITY (Chen *et al.*, 2010) was used for validation of the structure. Detailed statistics for the refined structure of hGal-7 complexes are given in tables 4.2 and 4.3. Figures were drawn with PYMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC). RMS calculations were performed using SUPERPOSE (Krissinel and Henrick, 2004). Ligand protein interactions were analysed using LIGPLOT+ (Laskowski and Swindells, 2011).

The atomic co-ordinates and structure factors (codes 4UW3, 4UW4, 4UW5 and 4UW6) have been deposited in the Protein Data Bank, Research Collaboratory for structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

4.2.5. Isothermal titration calorimetry:

Isothermal titration calorimetry (ITC) is a physical technique used to determine the thermodynamic parameters of interactions in solution. In this study ITC was employed ITC to determine binding affinity of dendrons to hGal-7.

Principle of ITC

Formation of unique spatial structures of biological macromolecules are in principle reversible, thermodynamically driven reactions. Thus, direct measurement of heat effects associated with intra- and inter-macromolecular processes can be performed using ITC. The reaction occurs in the sample pool whose temperature will be different from the reference pool, and this difference will be detected by an ITC instrument. Exothermic reaction activates negative

feedback of constant temperature power, while decalescence reaction activates positive feedback which can make the temperature constant.

Experimental design

Purified hGal-7 was dialysed overnight against 50 mM Na phosphate, 150 mM NaCl and 10 mM β-mercaptoethanol at 4 ^{o}C. The dialysed protein was then concentrated to several concentrations (1 mg/mL, 2 mg/mL, 5 mg/mL) to optimise the concentration of protein required to perform ITC experiments. Dendrons (D1 and D2), D-galactose (GAL) and N-acetyl galactosamine (GALNAc), were dissolved into the dialysis buffer at varying concentrations (10 mM, 20 mM, 25 mM). Both, the protein and ligand samples were degassed. The dendrons were placed in the injection syringe while the protein was placed in the nanocalorimeter cell of a Low Volume Nano ITC 60102 (TA Instruments). The complete reaction chamber was maintained at 15 ^{o}C. 190 µL of the protein was titrated with the dendrons (15 injections of 3 µL at an interval of 180 s). Control experiments were performed by injecting dendrons in the buffer. Commercial GAL and GALNAc, prepared as in the same manner as the dendrons, were used as standards for binding affinity comparison. Data were fitted with NanoAnalyze software (TA Instruments) according to standard procedures. Fitted data yielded the stoichiometry (n), the dissociation constant (K_d), the association constant (K_a), the enthalpy of binding (ΔH) and entropy, ΔS. The changes in free energy, ΔG) were calculated from the equation $\Delta G = \Delta H - T\Delta S$, where T is the absolute temperature. Two independent titrations were performed for each ligand tested.

4.2.6. Dynamic light scattering

Dynamic light scattering (DLS) is a technique used to determine the size distribution profile of proteins nucleic acids, and complexes, to detect the state of polydispersity in a macromolecular solution.

Theory

Particles move in a solution by Brownian motion. Particles of different size move in solution at different speeds, the larger the molecule, the slower the movement in solution (Lorber *et al.*, 2012). On exposure of the protein solution to monochromatic coherent laser, the particles in solution scatter a small fraction of the light. The intensity of this scattered light is recorded by the detector. Due to the Brownian motion of particles, the intensity of scattered light detected fluctuates with a constant change in distances between particles. This constant movement of particles causes the scattered light to undergo either constructive or destructive interference. The measured intensity fluctuations provides information about how fast the particles move in solution from which its size can be predicted. For example, smaller particles have higher intensity fluctuations than larger particles as they move faster in solution.

The size of the particle in solution (R_H) is calculated using the Stokes-Einstein equation.

$$R_H = \frac{kT}{6\pi\eta D} \qquad \text{.....(7)}$$

Where,
k represents the Boltzmann's constant; T represents temperature
H represents the viscosity; D is the diffusion constant

The Zetasizer software autocorrelates the intensities of the scattered light for the duration of the experiment. The time autocorrelation function of scattered light is used to determine the distribution of size across a molecule.

Experimental details

The final purified hGal-7 (0.5 mg/mL) was used for DLS measurements. Prior to adding dendrimers, the protein was filtered with 0.22 μm Millipore Millex ® (EMD Millipore, MA, USA) filter units to eliminate dust interferences.

Dendrons were prepared by dissolving in the final buffer used for hGal-7. The protein, hGal-7 was pre-incubated with the dendrons for 4 h at room temperature at varying protein: ligand ratios of 6:1, 3:1 and 1:2. The readings were taken in triplicate. DLS measurements were acquired with Zetasizer Nano S (Malvern Instruments, U.K.) which uses a 4 mW He-Ne laser with wavelength (λ_0) = 633 nm, and a scattering angle of 173^o. A low volume 12 µL quartz cuvette (QS 3.00 mm) was equilibrated by rinsing with the final buffer. 50 µL of the sample was loaded into the cuvette and the external surfaces of the cuvette were wiped with soft lens cleaning tissue to ensure no particles were present that could lead to false reading. The cuvette was placed in the sample holder and allowed to equilibrate to 16 ^{o}C for 10 min. A series of three scans were performed and subsequent intensity fluctuations were recorded which contain information about the time scale movement of the scattering particles. Autocorrelation curves were analysed using the software provided by the manufacturer.

4.3. Results

4.3.1. Cloning, expression and purification of hGal-7

hGal-7 was expressed in *E. coli* BL21 (DE3) cells and purified using IMAC [Figure 4.3] followed by size exclusion chromatography [Figure 4.4]. The 16 kDa purified protein [Figure 4.5] was then used to set up crystallisation screens in complex with the multivalent dendrons and for isothermal calorimetry experiments to study the binding affinity of hGal-7 to the dendrons.

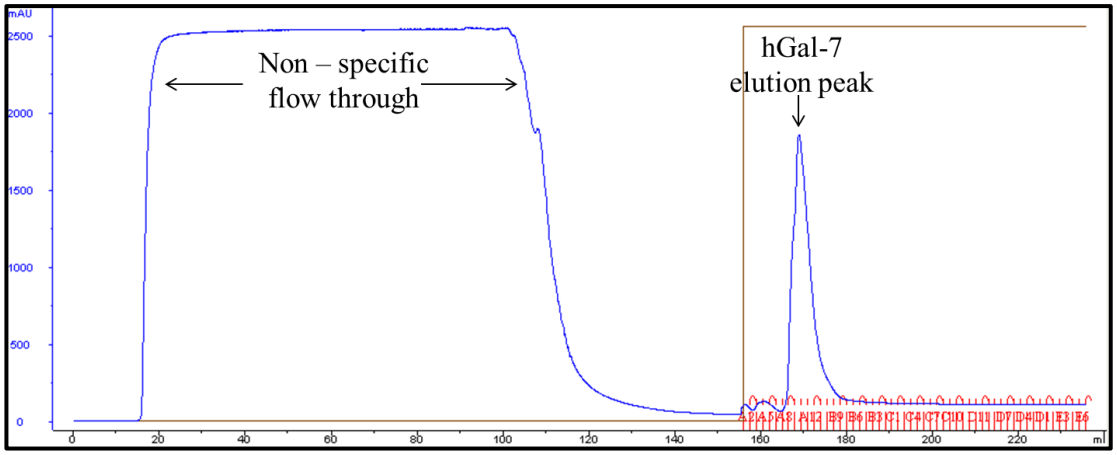


Figure 4. 3: hGal-7 purification step 1.

Example chromatogram of IMAC utilising the HisTrap FF column for purification of His-tagged hGal-7. Blue line represents the absorbance at A_{280} and the brown line represents the percentage of elution buffer.

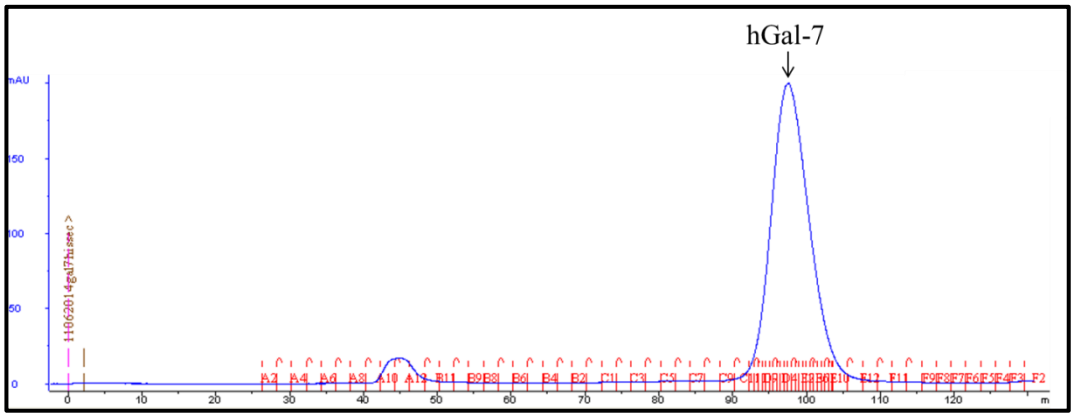


Figure 4. 4: hGal-7 purification step 2.

Example chromatogram of purification of hGal-7 by SEC using superdex 200 column. Blue line represents the absorbance at A_{280}.

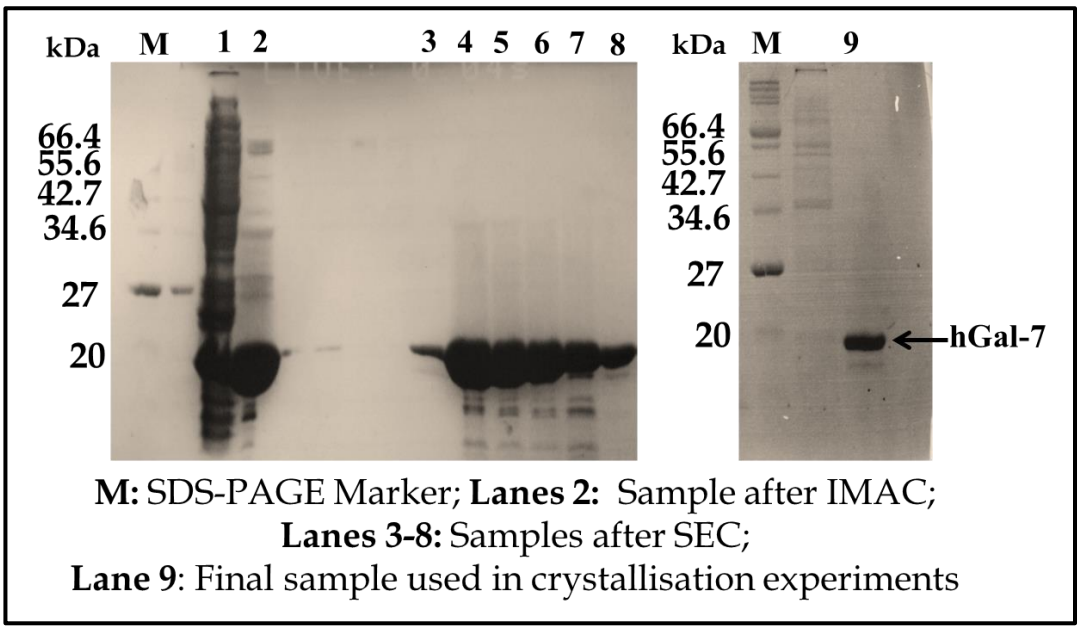


Figure 4. 5: Example of hGal-7 purification analysed by SDS-PAGE.
hGAl-7 purification analysed by SDS-PAGE using IMAC (lane 2) followed by SEC (lanes 3-8). The final purified sample is shown in lane 9. The molecular weights of the SDS-PAGE marker (M) are labelled.

4.3.2. Crystallisation of hGal-7 with dendrimers

Several crystallisation hits were observed for preliminary crystallisation screening of hGal-7 in complex with the dendremeric molecules. Of the 12 commercial screens, PACT *premier* HT-96, JCSG-*plus* HT-96, Morpheus HT-96 and PGA Screen HT-96 gave the most number of hits.

hGal-7 in complex with dendrons yielded crystallisation hits within 4 - 7 days of set up. The crystals were either small needles or microcrystalline powder which required further optimisation trials. It was observed that a protein to dendron ratio of 3:1 and an incubation temperature of 16 ^{o}C was favourable for crystal growth. Two different morphologies of the crystals for hGal-7 in complex with D2 were observed. They were either needle like crystals or small rod-like crystals.

It was also observed that poly(ethylene glycol) 3350 (PEG 3350) and BisTris propane pH 7 yielded more crystals than other conditions. To optimise the

protein concentration required for crystallisation, PACT *premier* HT-96 and JCSG-*plus* HT-96 screens were set up using the robot. Using these screens, protein concentration optimisation trials were set up with a protein concentration of 12mg/mL, 10 mg/mL, 8 mg/mL, 6 mg/mL and 4 mg/mL. The hits obtained from these screens were then used to set up manual optimisation trials. To reduce the rate of equilibration and to slow down the process of crystallisation, paraffin oil and silicone oil were added to the reservoir prior to sealing. This helped yield larger crystals that were suitable for X-ray data collection. The final conditions for crystallisation of hGal-7 in complex with the D1, D2 and D3 for which x-ray diffraction data were collected are listed below.

hGal-7 in complex with D1 was crystallised in 0.05 M BisTris Propane (pH 7), and 17% [w/v] PEG 3350, at a protein concentration of 12 mg/mL and protein to ligand ratio of 3:1. Paraffin oil was used to slow down the rate of equilibration.

hGal-7 in complex with D2 crystallised in two different conditions. Crystal form, D2-1 crystallised in 0.05 M BisTris Propane (pH 7), and 17% [w/v] PEG 3350, at a protein concentration of 12 mg/mL and protein to ligand ratio of 3:1. Crystal form, D2-2 crystallised in 0.05 M BisTris Propane (pH 7), and 14% [w/v] PEG 3350, at a protein concentration of 10 mg/mL and protein to ligand ratio of 3:1. Paraffin oil was used for D2-2 to slow down the rate of equilibration.

hGal-7 in complex with D3 was crystallised in 0.05 M BisTris Propane (pH 7), and 17% [w/v] PEG 3350, at a protein concentration of 6 mg/mL and protein to ligand ratio of 3:1. Paraffin oil was used to slow down the rate of equilibration.

Despite varying various conditions for crystallisation of hGal-7 in complex with dendrons D4, D5 and D6, no diffractable crystals were observed.

Figure 4.11, 4.12, 4.13 and 4.14 show the crystals shot for data collection at the Diamond Light Source (DLS).

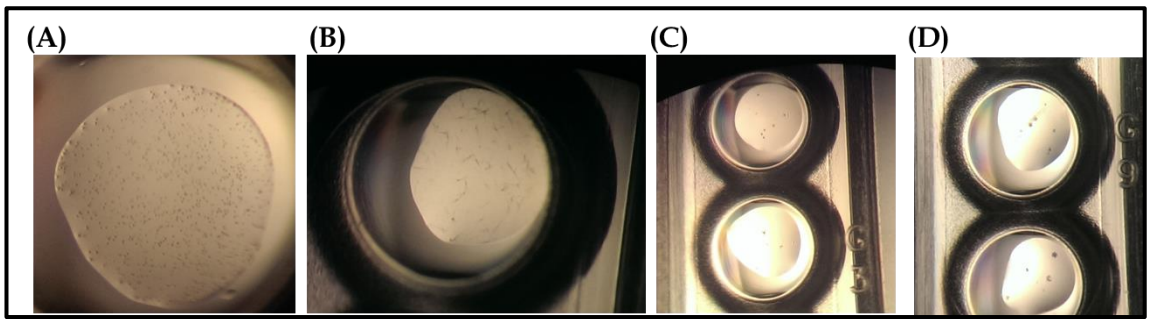


Figure 4. 6: Examples of initial hits observed for hGal-7 in complex with dendrons D1 and D3

(A-B) hGal-7 in complex with D1. (C-D) hGal-7 in complex with D3. Conditions: (A) C3 of JCSG-*plus* HT-96, (B) E6 of PACT *premier* HT-96, (C) F7 of PACT *premier* HT-96, (D)A10 of JCSG-*plus* HT-96.

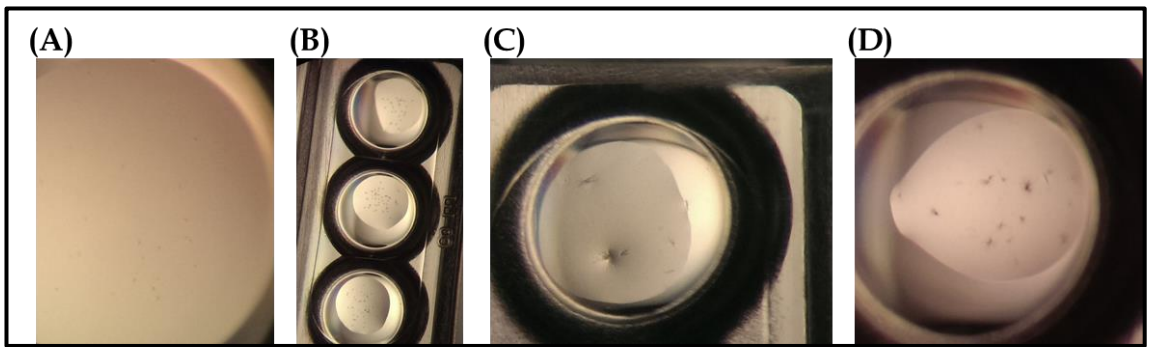


Figure 4. 7: Examples of initial hits observed for hGal-7 in complex with D2

(A-B) Crystal morphology 1 (C-D) Crystal morphology 2. Conditions: (A) C3 of JCSG-*plus* HT-96, (B) A8 of JCSG-*plus* HT-96, (C) G3 of PACT *premier* HT-96 (D) F2 of PACT *premier* HT-96.

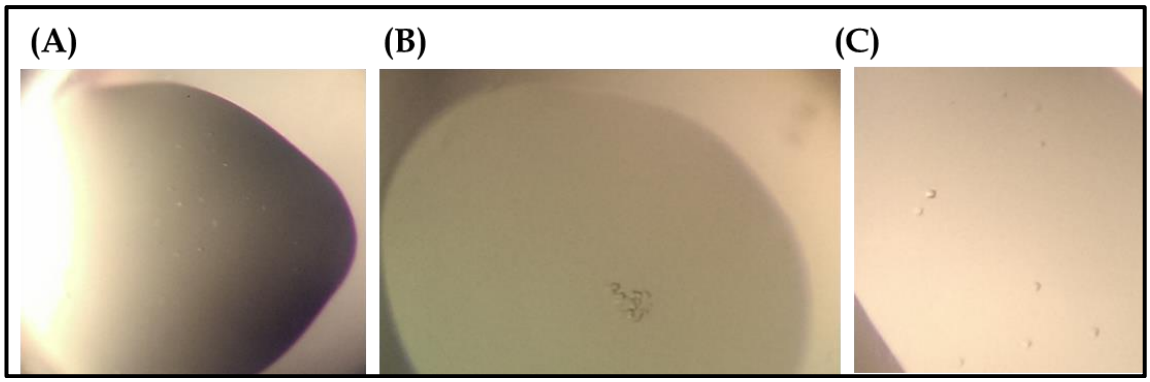


Figure 4. 8: Examples of initial hits observed. hGal-7 in complex with (A) D4 (B) D5 and (C) D6 with different morphology.

Conditions: (A) G3 of PACT *premier* HT-96, (B) G4 of PACT *premier* HT-96, (C) A9 of JCSG-*plus* HT-96

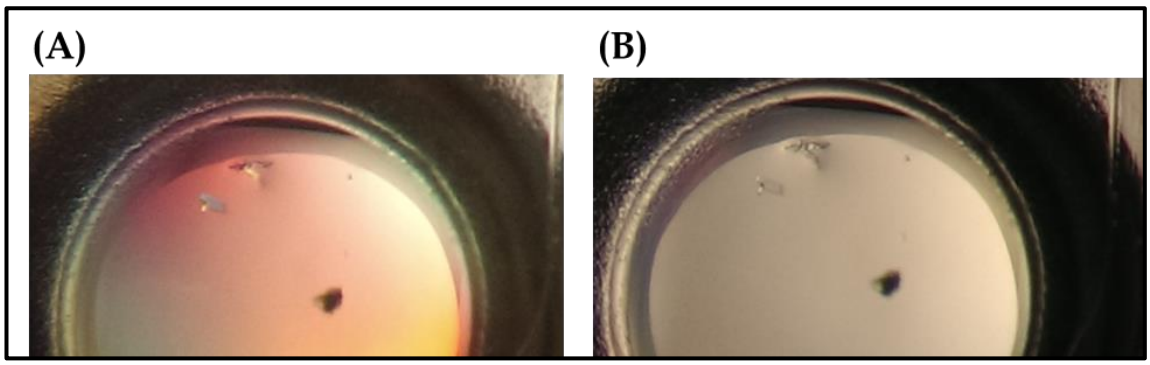


Figure 4. 9: Example of optimisation hits for hGal-7 in complex with D1.

(A) under polarised light (B) under white light.

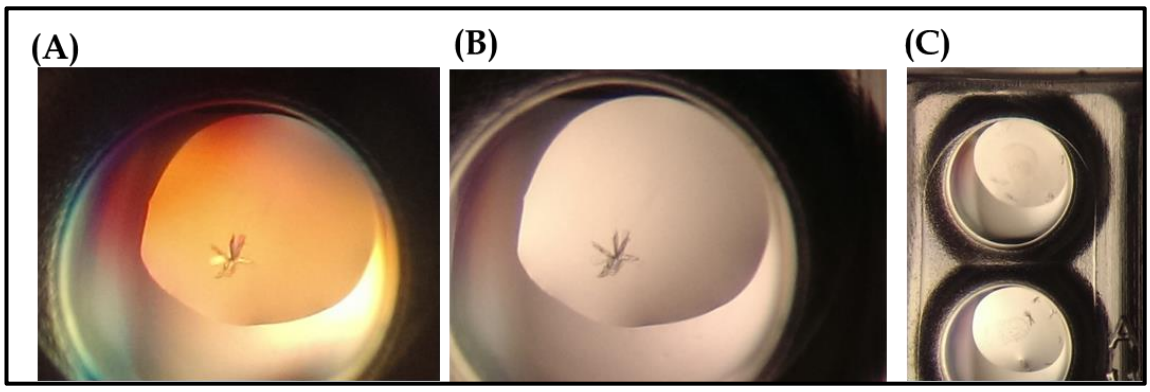


Figure 4. 10: Examples of optimisation hits for hGal-7 in complex with D2.

(A) morphology 1 observed under polarised light (B) Crystal morphology 1 observed under white light (C) Crystal morphology 2–no biorefringence observed.

4.3.3. Crystallographic results for hGal-7 in complex dendrimers.

Of the numerous crystals taken for data collection to the synchrotron, only a few diffracted to high resolution (>2.5Å). On processing these data sets, the structures revealed that on many occasions the proteins crystallised in their native form with no dendrimer bound at the CRD.

However, four datasets were collected for dendron bound hGal-7.

A complete dataset for hGal-7 in complex with D1 to 1.48 Å were collected on beam-line I04 from a crystal in the space group $P2_12_12_1$ with unit cell dimensions a = 30.4, b = 65.3, c = 114.1 Å and two molecules per asymmetric unit [Figure 4.11].

Data were collected from two forms of hGal-7 in complex with D2 in the space group $P2_12_12_1$. The first crystal form (D2-1) was collected on beam-line I04 to 1.86 Å with unit cell dimensions a = 38.91, b = 58.42, c = 114.84 Å and two molecules in the asymmetric unit [Figure 4.11]. The second crystal form (D2-2) was collected on the beam-line I03 to 2.0 Å with unit cell dimensions a = 63.4, b = 115.65, c = 116.25Å and six molecules per asymmetric unit [Figure 4.11].

Data from the co-crystal of hGal-7 in complex with D3 was collected on the beam-line I04-1 to 1.79 Å in the space group C2 with unit cell dimensions a = 58.93, b = 53.38, c = 89.21 Å, and β=102.8^o, and two molecules in the asymmetric unit [Figure 4.11].

Table 4.2 gives details of the data collection and refinement statistics. Figure 4.12 shows the representative electron density of hGal-7 in complex with the dendrons. Ramachandran plots are given in appendix D.

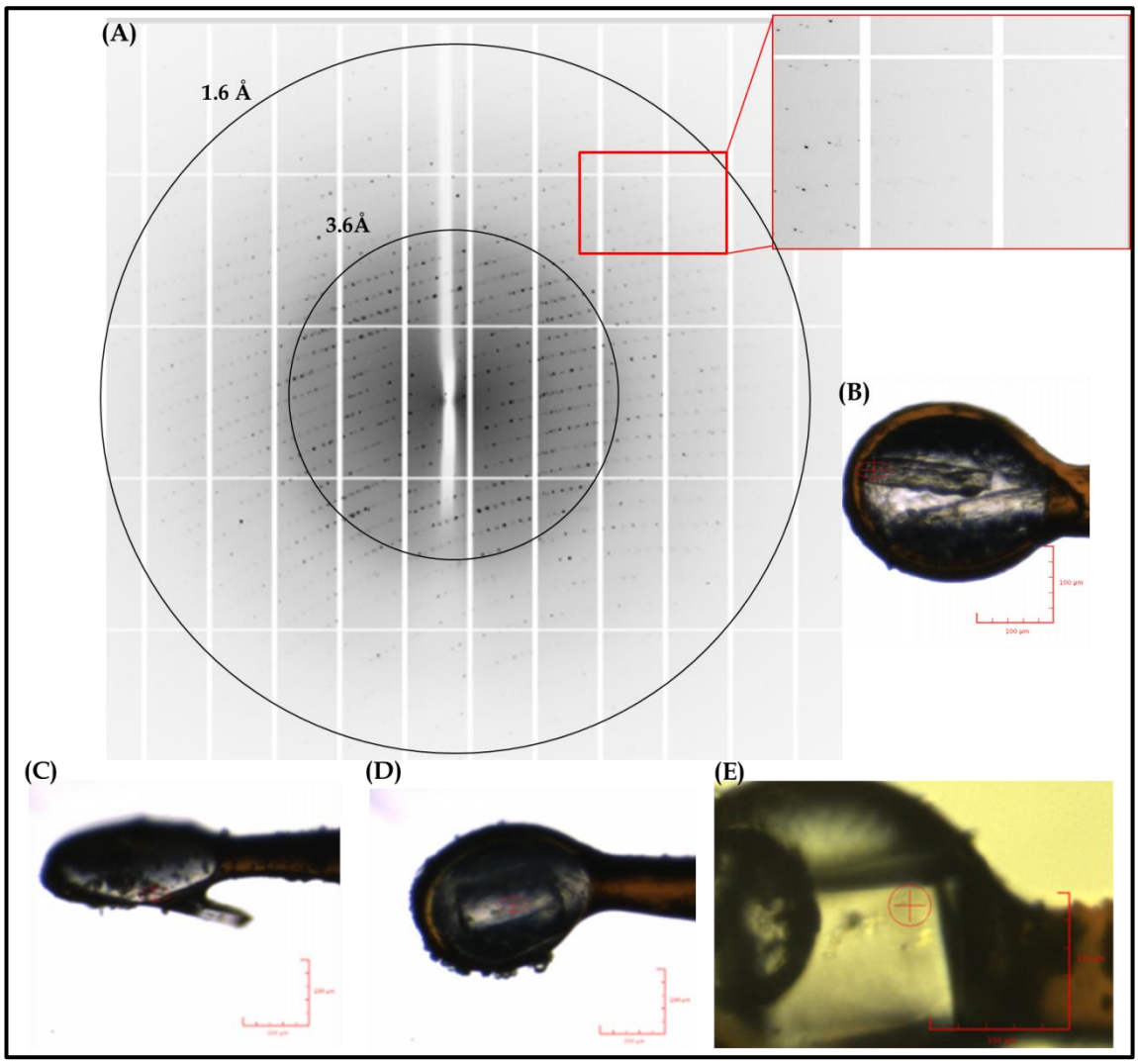


Figure 4. 11: Crystals and X-ray diffraction of hGal-7 in complex with dendrons.

(A) Example of the diffraction image collected at DLS beam-line I04 where the crystals of hGal-7 in complex with D1 diffracted to 1.48. Å. (B-E) Crystals mounted in a loop for data collection. (B) hGal-7 in complex with D1, (C) hGal-7 in complex with D2-1, (D) hGal-7 in complex with D2-2 and (E) hGal-7 in complex with D2-3.

Table 4. 2: X-ray data collection and refinement statistics for hGal-7 in complex with D1 and D2-1.

	hGal-7 – D1	**hGal-7– D2-1**
Data Collection Statistics		
Space group	$P2_12_12_1$	$P2_12_12_1$
Number of molecules / asymmetric unit	2	2
Cell dimensions	*a* = 30.4, *b* = 65.4, *c* = 114.1 Å	*a* = 39, *b* = 58.5, *c* = 115.9 Å
Resolution range (outer shell), Å	29.38 – 1.48 (1.56 – 1.48)	57.5 – 1.77 (1.83 – 1.77)
R_{symm}[a] (outer shell), %	7.5 (48.6)	5.3 (7.5)
I/σI (outer shell)	16.5 (4.3)	17.0 (9.5)
Completeness (outer shell), %	99.8 (98.34)	96.6 (91.5)
Redundancy (outer shell)	8.4 (6.9)	3.4 (2.4)
Total number of reflections	328115	87011
Unique no. of reflections	38926	25666
Wilson B-factor (Å^2)	13.59	12.72
Refinement Statistics		
R_{cryst}[b] (%)	17.6	18.2
R_{free}[c] (%)	19.8	23.6
Number of non-H atoms Protein Ligand Water molecules	2398 2100 42 256	2451 2118 64 269
Average B-factor (Å^2) Overall Protein Ligand Solvent	 17.6 16.4 22.5 27.0	 16.0 14.6 22.7 25.7
RMSD in Bond length (Å) Bond angle (°)	 0.007 1.18	 0.007 1.16
Ramachandran plot statistics Favoured (%) Allowed (%) Outliers (%)	 98.1 1.9 0	 98.9 1.1 0
RCSB PDB Code	4UW3	4UW4

[a] $R_{symm} = \Sigma h \Sigma i \, | I(h) - Ii(h) | / \Sigma h \Sigma i Ii(h)$, where $Ii(h)$ and $I(h)$ are the ith reflection and the mean measurements of the intensity of reflection h, respectively. [b] $R_{cryst} = \Sigma h \, | F_o - F_c | / \Sigma h F_o$ where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h, respectively. [c] R_{free} is equal to R_{cryst} for a randomly selected subset of reflections not used in the refinement.

Table 4. 3: X-ray data collection and refinement statistics for hGal-7 in complex with D2-2 and D3.

	hGal-7– D2-2	**hGal-7–D3**
Data Collection Statistics		
Space group	$P2_12_12_1$	C2
Number of molecules / asymmetric unit	6	2
Cell dimensions	*a* = 63.4, *b* = 115.6, *c* = 116.2 Å	*a* = 58.9, *b* = 53.4, *c* = 89.2 Å and β=102.9°
Resolution range (outer shell), Å	55.69 – 2.04 (2.11 – 2.04)	45.49 – 1.79 (1.85 – 1.79)
R_{symm}[a] (outer shell), %	7.3 (59.8)	5.6 (20.9)
I/σI (outer shell)	9.1 (2.1)	8.2 (3.4)
Completeness (outer shell), %	98.7 (99.3)	95.2 (91.5)
Redundancy (outer shell)	3.3 (3.3)	2.6 (2.5)
Total number of reflections	178525	62672
Unique no. of reflections	54509	24387
Wilson B-factor (Å^2)	32.06	17.93
Refinement Statistics		
R_{cryst}[b] (%)	23.6	20.0
R_{free}[c] (%)	31.5	25.1
Number of non-H atoms Protein Ligand Water molecules	6856 6240 184 432	2343 2092 95 156
Average B-factor (Å^2) Overall Protein Ligand Solvent	 35.8 35.8 35.9 36.0	 24.6 23.5 40.2 30.4
RMSD in Bond length (Å) Bond angle (°)	 0.008 1.17	 0.008 1.54
Ramachandran plot statistics Favoured (%) Allowed (%) Outliers (%)	 96 3.5 0.5	 98.8 1.2 0
RCSB PDB Code	4UW5	4UW6

[a] $R_{symm} = \Sigma h \Sigma i \, | I(h) - Ii(h) | / \Sigma h \Sigma i Ii(h)$, where $Ii(h)$ and $I(h)$ are the ith reflection and the mean measurements of the intensity of reflection h, respectively. [b] $R_{cryst} = \Sigma h \, | F_o - F_c | / \Sigma h F_o$ where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h, respectively. [c] R_{free} is equal to R_{cryst} for a randomly selected subset of reflections not used in the refinement.

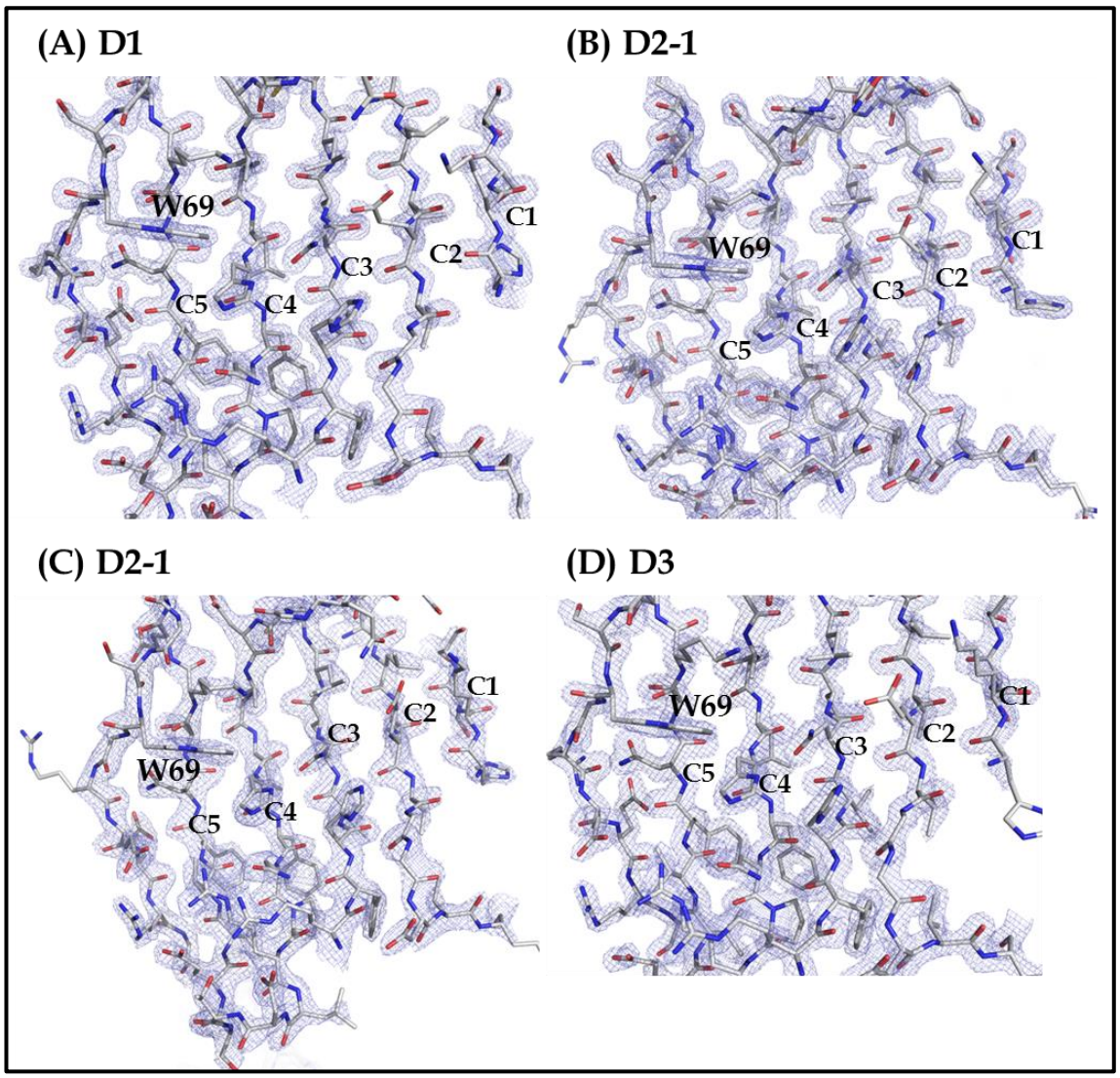


Figure 4. 12: Representative electron densities of hGal-7 in complex with dendrons.

The CRD is formed by residues of the β- sheets (C1-C5) and the adjoining loop residues. Residues are shown as sticks and the electron density maps contoured at 1σ.

4.3.2.1. Overall structure of hGal-7

The crystal structure of hGal-7 in complex with dendrons D1, D2 (in two crystal forms -1 and -2), and D3 were determined at 1.48, 1.86, 2.00 and 1.79 Å resolution, respectively. The overall structure of hGal-7 [Figure 4.13] remains unchanged upon ligand binding and appears as a dimer comparable to that previously described (Leonidas *et al.*, 1998). The 'jelly-roll' fold of galectin is completely observed in the electron density maps except for residues 1-3 at the N-terminal end.

Various parts of the carbohydrate ligands are visible in all of the structures with a common feature being at least one galactose moiety of each ligand bound at the CRD. Interaction between the protein and ligand resulted in cross-linking of hGal-7 thus forming a continuous sequence of 'dimer–ligand–dimer' repeats, with binding of the ligand also impacting on the hGal-7 dimeric interface.

The characteristic Π-Π stacking interaction between W69 and the terminal galactosyl ring is observed in the hGal-7 complex structure. The dendron binding is strengthened by additional hydrogen bonds and non-bonded interactions between various parts of the dendron and the CRD. Water molecules strengthen ligand binding through water mediated interactions. Ligand binding at the CRD also has a long range effect on the dimerisation interface. All these interactions have been discussed below.

Due to the subtle differences in the mode of ligand binding and dimerisation in each complex, the resultant monomers adopt significantly different crystal-packing contacts in each structure. This results in the formation of different supramolecular assemblies which have been analysed later.

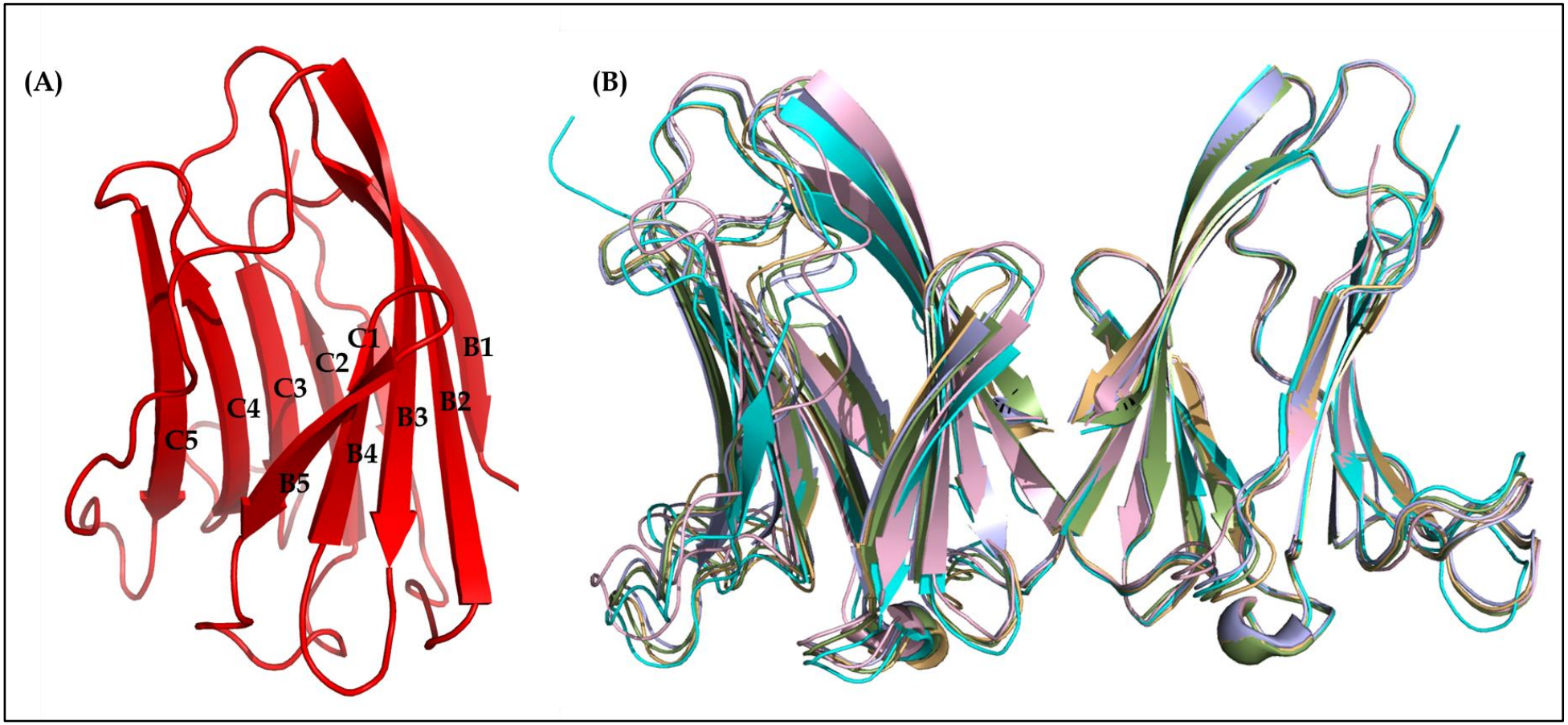


Figure 4. 13: Cartoon structure of hGal-7. (A) hGal-7 monomer, (B) Superposition of hGal-7 complexes.

(A) The β-strands forming the jelly-roll fold of hGal-7 have been labelled. (B) The structures of hGal-7 in complex with D1(light pink), D2-1(violet), D2-2(green), D3(light orange) and galactose (cyan) are superposed using the Cα atoms of Chain A (monomer on the right).

4.3.2.2. Dimeric Interface

The dimeric interface is based on the β-strand (B1-B5 interaction) [Figure 4.13] of two monomeric hGal-7 molecules and has been described extensively for the native structure (Leonidas *et al.*, 1998). The two monomers are associated by a two-fold rotational symmetry axis forming an extended β-sheet. Ligand binding does not appear to disrupt the dimeric state of hGal-7 but slightly alters the interface of dimerisation. PISA analysis (Krissinel and Henrick, 2007) of this dimeric interface shows ~795 $Å^2$ interface area in the native structure (PDB: 3ZXF) (Masuyer *et al.*, 2012) associated with 14 hydrogen bonds and 21 salt bridges.

Ligand binding at the CRD results in a long range impact at the dimeric interface which has been analysed by calculating the RMS deviation for the main chain, C^{α} and the side chains of the A and B chains of hGal7 in complex with each ligand with respect to the native hGal-7 chain A [Table 4.4]. Chain B in each complex shows a higher RMS deviation for the main chain and Cα atoms than chain A. However the RMS deviation for the side chains is comparably higher for chains A and B. Table 4.5 shows a comparison of the RMS deviation of the side chains involved in the dimerisation of hGal7 in each complex. Figure 4.14 shows the dimerisation surface of hGal-7 and Figure 4.15 shows the residues involved in the dimerisation of hGal-7.

Table 4. 4: RMS deviations for chains A and B in the complexes with respect to native hGal-7.

RMS Deviation (Å)	Galactose (PDB: 2GAL)	hGal-7 – D1	hGal-7 – D2-1	hGal-7 – D2-2	hGal-7 –D3
Main Chain -					
Chain A	0.54	0.74	0.60	0.44	0.63
Chain B	0.77	0.82	1.10	1.07	1.15
Cα -					
Chain A	0.50	0.75	058	0.44	0.53
Chain B	0.79	0.82	1.11	1.07	1.13
Side Chain -					
Chain A	1.18	1.61	1.48	1.47	1.65
Chain B	1.66	1.66	1.90	1.90	1.87

Table 4. 5: RMS deviations of side chain residues involved in the dimerisation of hGal7.

RMS Deviation (Å)– Side chain	Galactose		hGal-7 – D1		hGal-7 – D2-1		hGal-7 – D2-2		hGal-7 –D3	
Residue	**Chain A**	**Chain B**	**Chain A**	**Chain B**	**Chain A**	**Chain B**	**Chain A**	**Chain B**	**Chain A**	**Chain B**
R14	1.21	2.59	3.53	4.36	2.64	2.00	1.94	2.26	1.55	2.88
T17	0.33	0.28	0.51	0.20	0.51	0.48	0.79	0.51	0.56	0.28
R20	0.47	1.46	1.45	0.53	0.51	0.62	2.17	1.55	2.21	1.60
R22	0.53	1.33	0.39	1.39	0.30	1.63	0.65	1.87	0.36	1.25
E87	0.46	1.53	0.50	0.86	0.36	1.61	0.64	1.61	0.41	1.57
S93	1.93	0.26	0.23	0.54	0.07	0.32	0.41	1.12	0.23	0.60
D94	0.86	0.59	0.46	3.02	2.92	0.39	2.96	1.25	0.33	2.35
D95	2.55	0.98	1.60	0.46	1.01	1.46	1.45	1.28	2.58	1.27
K98	0.61	1.02	2.01	0.81	0.88	1.08	0.92	1.49	0.58	0.86
F135	0.81	0.76	0.42	0.86	1.22	0.91	1.11	0.83	1.12	1.00

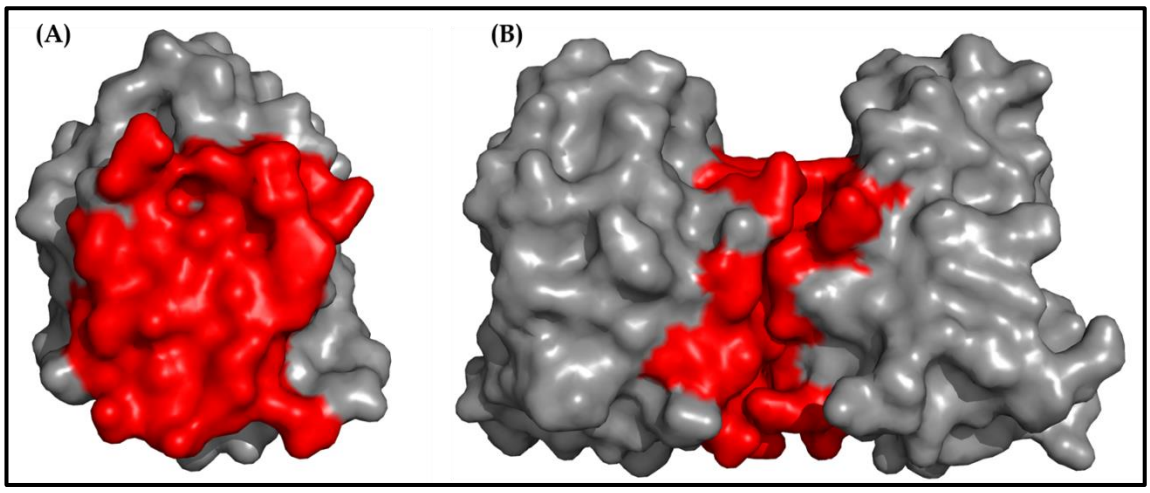


Figure 4. 14: Surface of the dimeric interface of hGal-7.

A) Surface of the monomer that interacts with the second molecule of hGal-7 to form the homodimer. (B) The hGal-7 dimer. The interacting surfaces are shown in red.

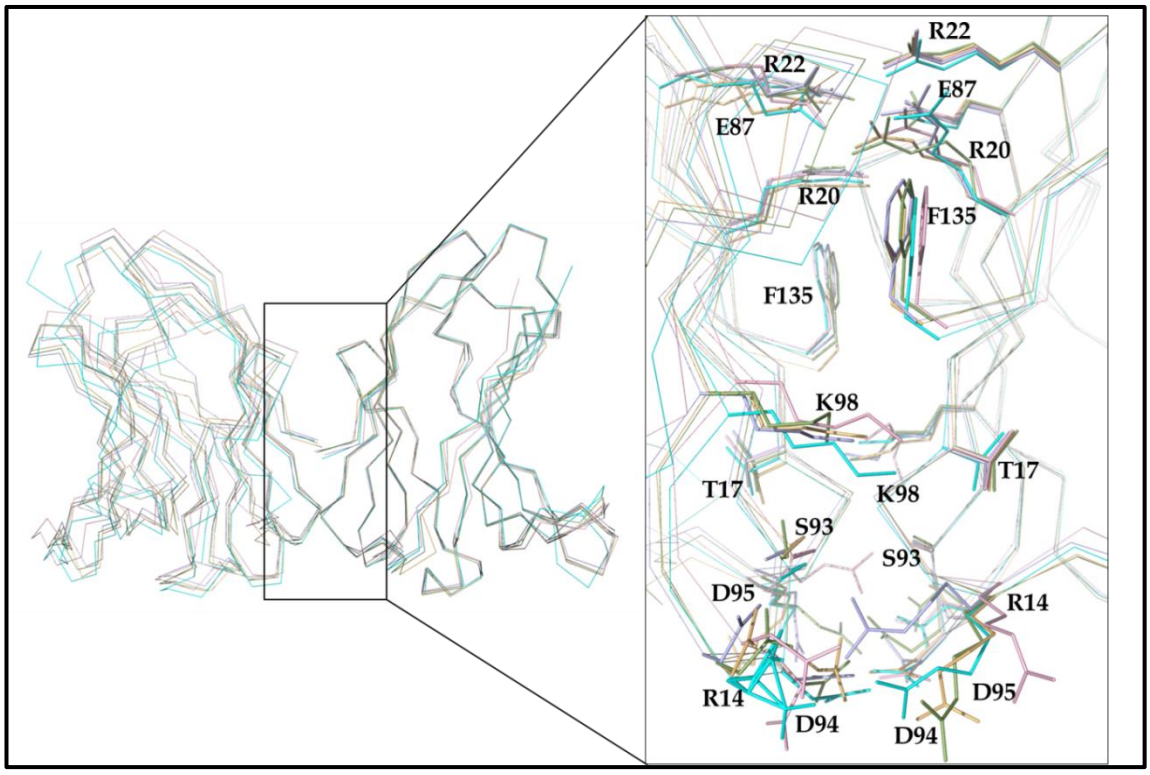


Figure 4. 15: The dimeric interface of hGal-7 in the crystal structure.

The structures of hGal-7 in complex with D1 (light pink), D2-1 (violet), D2-2 (smudge), D3 (light orange) and galactose (cyan) are superposed using the Cα atoms of Chain A. The inset highlights the residues (shown as sticks) that are involved in hydrogen bond formation at the dimeric interface.

PISA analysis of the hGal-7 dimer in complex with D1 shows a dimeric interface of 758 Å^2 associated with 13 hydrogen bonds and 14 salt bridges (Krissinel and Henrick, 2007). The hydrogen bonds between residues R14-D95, R20-D103 and D103-R133, which are observed in the native hGal-7 structures, dissociate in this complex and additional bonds between residues R22-D103 and R14-S93 are formed [Table 4.6]. Salt bridges between R14-D95, R20-D103, and D103-R133 dissociate while an additional salt bridge between R22–D103 is observed [Table 4.7].

hGal-7 in complex with D2-1 shows a dimeric interface of 774 Å^2 associated with 17 hydrogen bonds and 16 salt bridges on PISA analysis (Krissinel and Henrick, 2007). As compared to the dimeric interface of the native dimer, most hydrogen bonds are retained in this the hGal-7 complex with D2-1, with the exception of those between R14-D95 and D103-R133 [Table 4.6]. The salt bridges between R20-E87 and R133-D103 that are observed in the native hGal-7 structure, dissociate at the dimeric interface in this complex [Table 4.7]. However, an overall decrease in the bond distance between residues forming the non-bonded contacts at the dimeric interface is observed.

hGal-7 in complex with D2-2 shows a dimeric interface of 772 Å^2 associated with 16 hydrogen bonds and 20 salt bridges. As compared to native hGal-7, additional hydrogen bonds between R14-D94, R22-E87 and D103-R133 are formed in this complex [Table 4.6]. Further, the non-bonded contacts between D103-R133 dissociate, while between R14-D94 they are formed [Table 4.7].

The crystal structure of hGal-7 in complex with D3 shows a dimeric interface of 766 Å^2 associated with 15 hydrogen bonds and 16 salt bridges. Hydrogen bonds at the dimeric interface between R22-D103 and R14-D94 are formed additional to those observed in the native structure [Table 4.6]. It is also observed that the hydrogen bonds between R14-D95, R22-E87 and R133-D103 dissociate in the structure of this complex. The non-bonded contacts at the dimeric interface between R14-D95, R20-E87, R20-D103, R22-E87 and R133-D103 dissociate in this

structure and additional hydrogen bonds between R14-D94 and R22-D103 are formed [Table 4.7].

Table 4. 6: Hydrogen bond distances between residues forming the dimeric interface in the complexes of hGal-7 with the dendrons.

All contacts shorter than 3.3 Å are listed.

		Distances (Å)				
Residue(A)	**Residue (B)**	**hGal-7-D1**	**hGal-7-D2-1**	**hGal-7-D2-2**	**hGal-7-D3**	**Native hGal-7**
S93 – OG	R14 – NH2	3.2	-	-	-	-
D94 – OD2	R14 – NH2	-	-	2.2	-	-
D95 – OD1	R14 – NH2	-	-	2.9	-	-
G102 – O	R20 – NH2	2.8	2.8	3.1	2.9	3.1
D103 – OD1	R20 – NH2	-	-	3.0	-	2.7
E87 – OE1	R22 – NH2	2.8	2.5	2.5	-	-
D103 – OD1	R22 – NH1	2.7	-	-	2.2	-
R20 – NH1	E87 – OE2	3.0	-	-	-	-
R22 – NH2	E87 – OE1	3.3	2.7	2.8	-	3.4
R14 – NH1	D94 – OD2	-	3.3	-	2.7	-
R14 – NH2	D95 – OD1	-	-	-	-	2.3
F135 – O	K98 – NZ	3.3	-	-	3.1	2.9
F135 – OXT	K98 – NZ	-	2.8	-	2.6	-
R20 – NH2	G102 – O	2.7	-	3.2	2.8	3.0
R20 – NH2	D103 – OD1	-	2.7	-	2.9	2.8
R22 – NH1	D103 – OD1	3.2	-	-	2.5	-
R133 – NE	D103 – O	2.9	2.8	-	-	-
R133 – NH1	D103 – O	-	-	2.7	-	-
D103 – O	R133 – NH1	-	-	-	-	2.7
D103 - OD2	R133 – NH2	-	-	-	-	2.9
D103 – O	R133 – NE	-	3.0	2.8	-	-
K98 – NZ	F135 – O	-	-	-	2.9	-
K98 – NZ	F135 – OXT	-	2.9	3.1	2.9	2.8

Table 4. 7: Distances between the residues involved in the formation of salt bridges at the dimeric interface.

Contact distances with the maximum allowed values of C-C, 4.1 Å ; C-N, 3.8 Å, C-O, 3.7 Å; O-O, 3.3 Å; O-N, 3.4 Å; and N-N, 3.4 Å are listed.

		Distances (Å)				
Residue(A)	**Residue (B)**	**hGal-7-D1**	**hGal-7-D2-1**	**hGal-7-D2-2**	**hGal-7-D3**	**Native hGal-7**
D94 – OD2	R14 – NE	-	-	3.2	-	-
D94 – OD2	R14 – NH2	-	-	2.2	-	-
D95 – OD1	R14 – NH2	-	-	2.9	-	-
E87 – OE1	R20 – NH1	-	-	-	-	3.4
D103 – OD1	R20 – NH2	-	-	3.0		2.7
E87 – OE1	R22 – NH2	2.8	2.5	2.5	-	-
E87 – OE2	R22 – NH2	-	-	3.1	-	-
D103 – OD1	R22 – NH2	-	-	-	3.8	-
D103 – OD1	R22 – NH1	2.7	-	-	2.2	-
R20 – NH1	E87 – OE2	3.0	-	3.3	-	-
R22 – NH2	E87 – OE1	3.3	2.7	2.8	-	-
R22 – NH2	E87 – OE2	-	3.3	-	-	3.4
R14 – NH1	D94 – OD2	-	-	-	2.7	-
R14 – NH1	D95 – OD1	-	3.3	-	-	-
R14 – NH2	D95 – OD2	-	-	-	-	2.3
F135 – O	K98 – NZ	3.3	-	-	3.1	2.9
F135 – OXT	K98 – NZ	-	2.8	-	2.6	-
R20 – NH1	D103 – OD1	-	-	-	-	3.7
R20 – NH2	D103 – OD1	-	2.7	-	2.9	2.8
R22 – NH1	D103 – OD1	3.2	-	-	2.5	-
R133 – NH2	D103 – OD2	-	-	-	3.3	-
D103 – OD2	R133 – NH2	-	-	-	-	2.9
K98 – NZ	F135 – O	-	3.4	3.4	2.9	-
K98 – NZ	F135 – OXT	-	2.9	3.1	2.9	2.8

4.3.2.3. Complex formation of hGal-7 with dendrons:

The CRD is highly conserved in hGal-7 (Leonidas *et al.*, 1998) and consists of residues H49, N51, R53, N62, W69, E72 and R74 on C4, C5 and the adjoining loop regions, which form a concave surface on the exposed surfaces of the dimer. The characteristic Π-Π stacking interaction of the galactosyl group with W69 is observed in all complexes. Additional hydrogen bonds and non-bonded contacts between the galactosyl termini of the dendrons and the residues at the CRD are also observed. Water molecules play an important role in the binding of dendrons to hGal-7 at the CRD. In addition to direct hydrogen bonding between the residues at the CRD and the ligands, a number of water molecules facilitate these interactions.

The electron density omit maps for the ligands in each complex are presented in figure 4.16. Electron density for two of the three arms for D1, two of the three arms for D2 (D2-1), all three arms of D2 (D2-2) and four of the six arms for D3 are observed in their respective complex structures with hGal-7.

The superposition of the terminal galactosyl ring of each of the dendrons with galactose is presented in figure 4.17. With the varying size, length and flexibility of the dendrons, the spatial orientation of the ligands in each complex is different. A comparison of this change in orientation is presented in figure 4.17.

Table 4.8 and 4.9 give details of the hydrogen bonds and non-bonded contacts formed between the residues at the CRD and the dendrons.

.

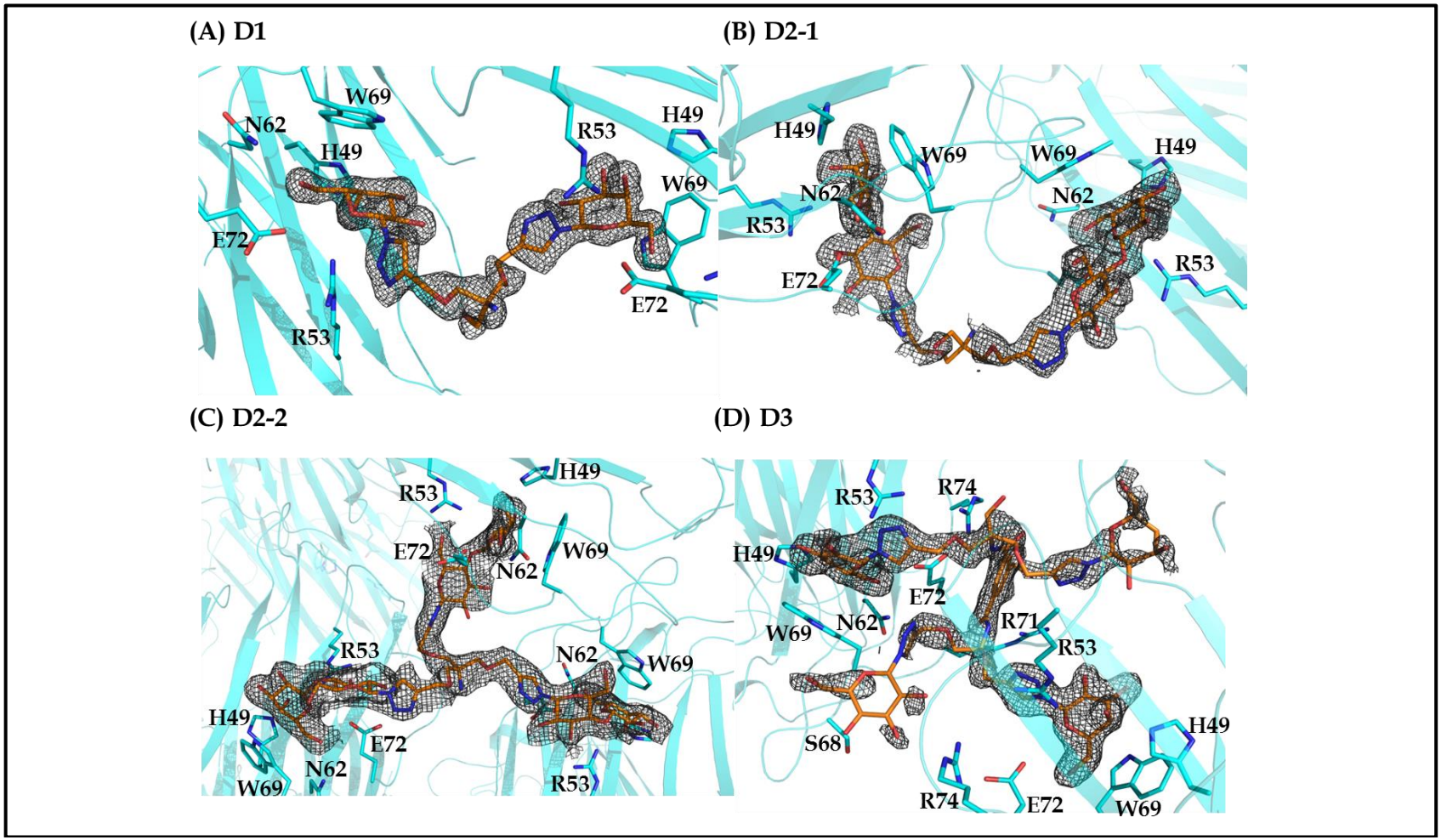


Figure 4. 16: Omit maps for hGal-7 in complex with dendrons.

(A) hGal-7 in complex with D1 (B) hGal-7 in complex with D2-1 (C) hGal-7 in complex with D2-2 and (D) hGal-7 in complex with D3. The omit maps are contoured at 1σ. Residues involved in the recognition and binding of the ligands are shown as sticks.

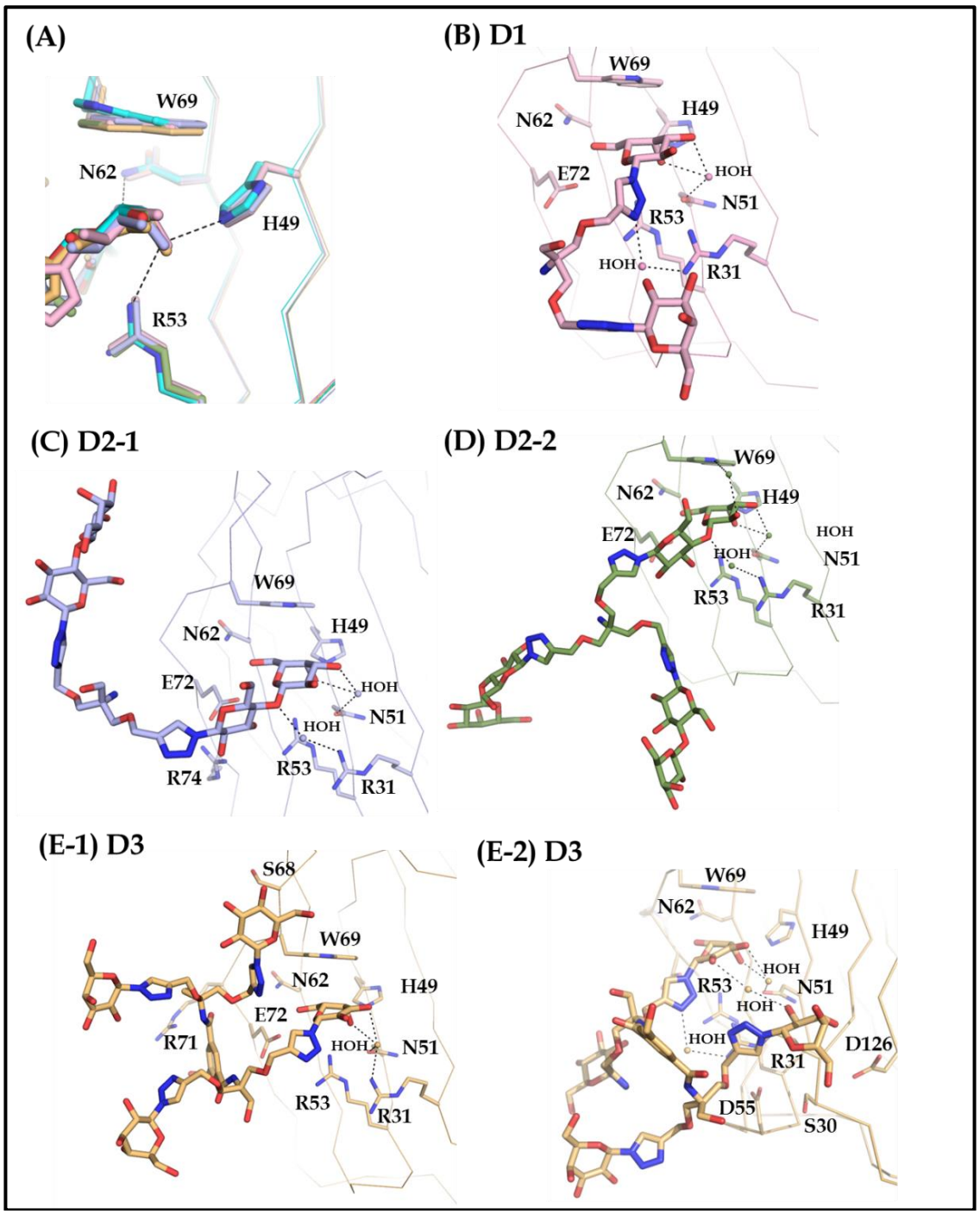


Figure 4. 17: CRD for hGal-7 in complex with dendrons.

(A) Superpose of hGal-7 in complex with D1 (light pink), D2-1 (pale cyan), D2-2 (smudge), D3 (light orange) and GAL (cyan). (B-E) All complex structures are shown in a similar spacial orientation. hGal-7 in complex with (B) D1 (C) D2-1 (D) D2-2 (E-1 and E-2) D3.E-1 and E-2 are symmetry related ligands bound to two different CRDs of the dimer. Residues involved in the recognition and binding of the ligands are shown as sticks. Water molecules that contribute to the bonding are shown as spheres.

Table 4. 8: Hydrogen bond distances between the ligand atoms and the atoms of the residues at the CRD in the complex structures.

All contacts shorter than 3.3 Å are listed.

Distances (Å)	hGal-7 -D1	hGal-7 -D2-1	hGal-7 -D2-2	hGal-7 -D3	hGal-7 - GAL
OED - OG S30	-	-	-	3.3	-
O8 – NE2 H49	2.7	-	-	-	-
O4 – NE2 H49	-	-	-	2.9	-
OCV – NE2 H49	-	-	2.9	-	-
OW3 – NE2 H49	2.7	2.9	2.8	2.9	2.7
OW1 – NH2 R53	2.9	3.0	3.1	3.1	-
OW2 – NH2 R53	3.0	-	-	-	-
OW3 – NH2 R53	2.9	2.9	2.9	2.9	3.2
OX4 – NH2 R53	-	3.0	2.9	-	-
OX4 – NH1 R53	-	2.9	2.6	-	-
O8 – NH2 R53	2.9	-	-	-	-
O5 – NH2 R53	3.0	-	-	2.9	-
OCV – NH2 R53	-	-	2.8	-	-
OCM – NH2 R53	-	-	3.1	-	-
ODG – NH2 R53	-	-	3.2	-	-
ODC – NH1 R53	-	-	2.4	-	-
O4 – NH2 R53	-	-	-	2.9	-
OCU – ND2 N62	-	-	2.8	-	-
O12 – ND2 N62	3.0	-	-	-	-
OW2 – ND2 N62	-	2.9	2.8	3.1	2.6
O6 – ND2 N62	-	-	-	2.9	-
OCF – NH2 R71	-	-	3.2	-	-
OCF – NE R71	-	-	3.1	-	-
ODG – NH1 R71	-	-	-	3.1	-
OZ3 – N E72	-	-	-	3.2	-
OX3 – NH2 R74	-	3.2	-	-	-
ODH – NH2 R74	-	-	3.2	-	-

Table 4. 9: Distances between the atoms of the residues at the CRD forming non-bonded interactions with the ligand in the complexes.

Contact distances with the maximum allowed values of C-C, 4.1 Å ; C-N, 3.8 Å, C-O, 3.7 Å; O-O, 3.3 Å; O-N, 3.4 Å; and N-N, 3.4 Å are listed.

Distances (Å)	hGal-7 -D1	hGal-7 -D2-1	hGal-7 -D2-2	hGal-7 -D3	hGal-7 - GAL
CAS - CB S30	-	-	-	3.8	-
OED - CB S30	-	-	-	3.6	-
CEV - NH2 R31	-	-	-	3.4	-
CEV - CZ R31	-	-	-	3.5	-
CBO - CZ R31	-	-	-	3.8	-
CEV - NE R31	-	-	-	3.8	-
CBO - NE R31	-	-	-	3.8	-
CAT - CD R31	-	-	-	3.8	-
CAC - CD R31	-	-	-	3.9	-
OAU - CD R31	-	-	-	3.2	-
CAC - CG R31	-	-	-	3.6	-
OAU - CG R21	-	-	-	3.4	-
CW3 – CD2 H49	3.9	-	-	-	3.8
CW3 – NE2 H49	-	3.5	3.3	3.7	3.3
CW1 – CD2 H49	-	3.8	3.8	3.7	3.6
OW3 – CD2 H49	3.5	-	3.6	3.4	3.5
OW3 – CE1 H49	3.7	-	-	-	-
OW3 – NE2 H49	3.3	-	-	-	-
NX6 – CZ R53	3.4	-	-	-	-
CX4 – NH1 R53	-	3.7	-	-	-
OX4 – CZ R53	-	3.4	3.2	-	-
CW2 – NH2 R53	3.7	3.8	-	-	-
CW1 – NH2 R53	3.7	-	-	-	-
CW5 – NH2 R53	-	3.7	3.8	-	-
CW6 – NH2 R53	3.7	3.6	3.7	-	-
CEV - OD2 D55	-	-	-	3.2	-
CAT - OD2 D55	-	-	-	3.6	-
OW2 – CG N62	-	-	-	-	3.5
OW2 – CB N62	-	-	-	-	3.6
CW1 – ND2 N62	-	-	3.7	-	3.7
CW1 – CB N62	-	-	3.9	-	-
CL1 – O G67	-	-	-	3.2	-
CL1 – CA S68	-	-	-	3.6	-
CL3 – CA S68	-	-	-	3.8	-
CL3 – CB S68	-	-	-	3.5	-
OL3 – CB S68	-	-	-	3.6	-
CW1 – CD2 W69	-	3.9	-	3.9	-
CW2 – CD2 W69	3.8	-	3.8	3.8	-
CW2 – CE2 W69	3.7	-	3.7	3.7	-
CW3 – CE2 W69	3.7	-	3.7	3.8	-
CW3 – CZ2 W69	3.6	-	3.6	3.7	3.8

CW4 – CZ2 W69	-	-	3.7	-	-
CW4 – CE2 W69	-	-	3.8	-	-
CM9 – CB W69	-	-	-	3.7	-
OL1 – CD1 W69	-	-	-	3.5	-
CW3 – CH2 W69	-	-	-	-	3.9
CW3 – CZ2 W69	-	-	-	-	3.8
CW1 – CE3 W69	-	-	-	-	3.9
CL1 – CD1 W69	-	-	-	3.8	-
CCP - CZ2 W69	-	3.8	-	-	-
CCO - CZ2 W69	-	3.7	-	-	-
CCP - CE2 W69	-	3.9	-	-	-
CCO - CE2 W69	-	3.7	-	-	-
CCN - CE2 W69	-	3.7	-	-	-
CN2 – O G70	-	-	-	3.6	-
CB2 – O G70	-	-	-	3.5	-
OZ3 – CD R71	-	-	-	3.4	-
ODG – CD R71	-	-	-	3.2	-
CA3 – CZ R71	-	-	-	3.8	-
CA3 – NH1 R71	-	-	-	3.2	-
CZ1 – NH1 R71	-	-	-	3.4	-
CZ3 – NH1 R71	-	-	-	3.0	-
CDF – NH1 R71	-	-	-	3.6	-
CBY - NH2 R71	-	-	3.8	-	-
OW2 – CG E72	3.4	3.3	3.1	3.4	3.2
OW2 – CD E72	3.4	3.4	3.2	3.3	3.5
CW1 – OE2 E72	-	3.6	3.5	-	3.7
CW8 – OE2 E72	-	3.6	-	-	-
CX4 – OE2 E72	-	3.3	3.1	-	-
CX3 – OE2 E72	-	3.6	-	-	-
CX3 – OE1 E72	-	3.6	-	-	-
OX4 – CD E72	-	3.2	-	-	-
CY6 – OE1 E72	-	-	-	3.6	-
CZ2 – OE1 E72	-	-	-	3.7	-
CA3 – CB E72	-	-	-	3.7	-
CDU – NH2 R74	-	-	-	3.6	-
CY9 – NH1 R74	-	-	-	3.7	-
CY8 – NH1 R74	-	-	-	3.5	-
CY7 – NH1 R74	-	-	-	3.3	-
CY7 – CZ R74	-	-	-	3.9	-
CAS - OD2 D126	-	-	-	3.0	-
CAS - OD1 D126	-	-	-	3.5	-
CAS - CG D126	-	-	-	3.6	-

The terminal monosaccharide (galactosyl ring) of D1 forms Π-Π stacking interactions with W69. Three hydrogen bonds between R53 and OW1, OW2 and OW3 of D1 and a hydrogen bond between H49 and OW3 of D1 contribute to the binding of the galactosyl ring [Table 4.8]. Residues N51 and R31 interact with OW4 and NX7 of D1, respectively, through water molecules [Figure 4.17B, 4.18]. Non-bonded contacts between H49, R53, W69 and E72 with D1 are also observed [Table 4.9]. No significant changes in the ligand binding are observed between chain A and chain B of the hGal-7 dimer [Figure 4.18].

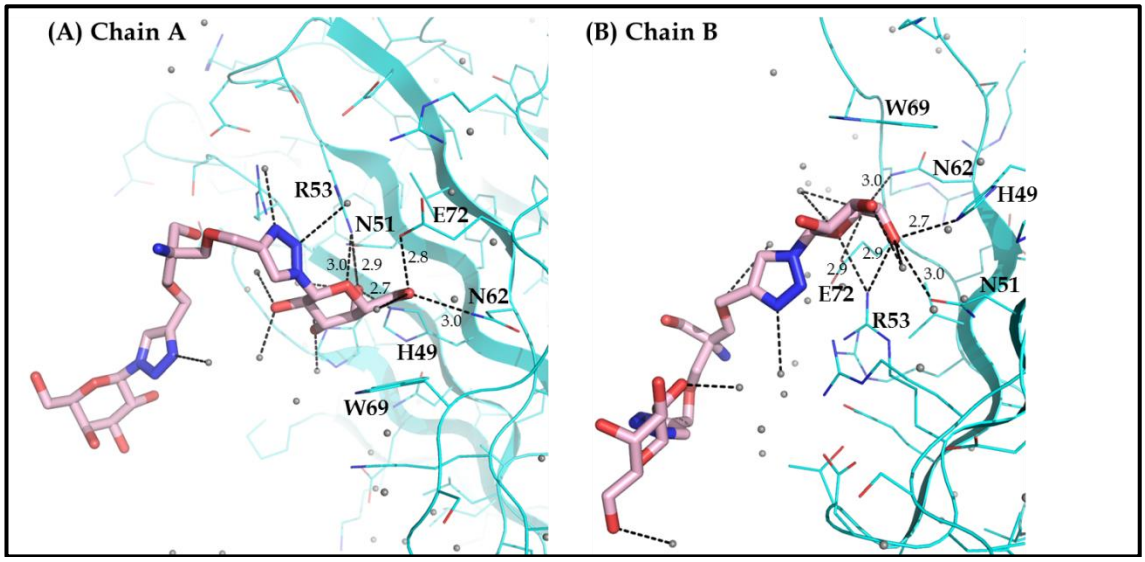


Figure 4. 18: CRD of hGal-7 in complex with D1.

Hydrogen bond interactions with both the chains of the dimer are shown. D1 is shown as sticks in light pink. Water molecules are shown as grey spheres. Distances are given in Å.

The terminal monosaccharide (galactosyl ring) of D2 forms Π-Π stacking interactions with W69 and hydrogen bonds with other amino acids of the CRD. Additional hydrogen bonds and other non-bonded contacts between the penultimate galactosyl ring and the CRD strengthens dendrimer recognition and interaction with hGal-7.

In the first crystal form, D2-1, residues H49, R53, N62 and R74 are involved in hydrogen bonding with the ligand [Table 4.8]. Most of these bonds are observed in the second crystal form, in addition to interaction with R71 [Table 4.8]. Both crystal forms show non-bonded interactions between residues H49,

R53, W69 and E72 with the ligands [Table 4.9]. In addition to this, hGal-7 in complex with D2-2 interacts with N62 and R71 through non-bonded contacts. The interactions between the ligand and residues N51, R31, W69 that are mediated through a water molecule are observed in both crystal forms [Fig. 4.17C, 4.17D, 4.19, 4.20].

In hGal-7 complexed with D2-1, the hydrogen bond observed between D2-1 and N51 at the CRD of chain A is not observed at the CRD of chain B [Figure 4.22]. This could be due to subtle changes at the CRD observed in both the chains. No significant changes in the ligand binding are observed between chain A and chain B of the hGal-7 dimer [Figure 4.23].

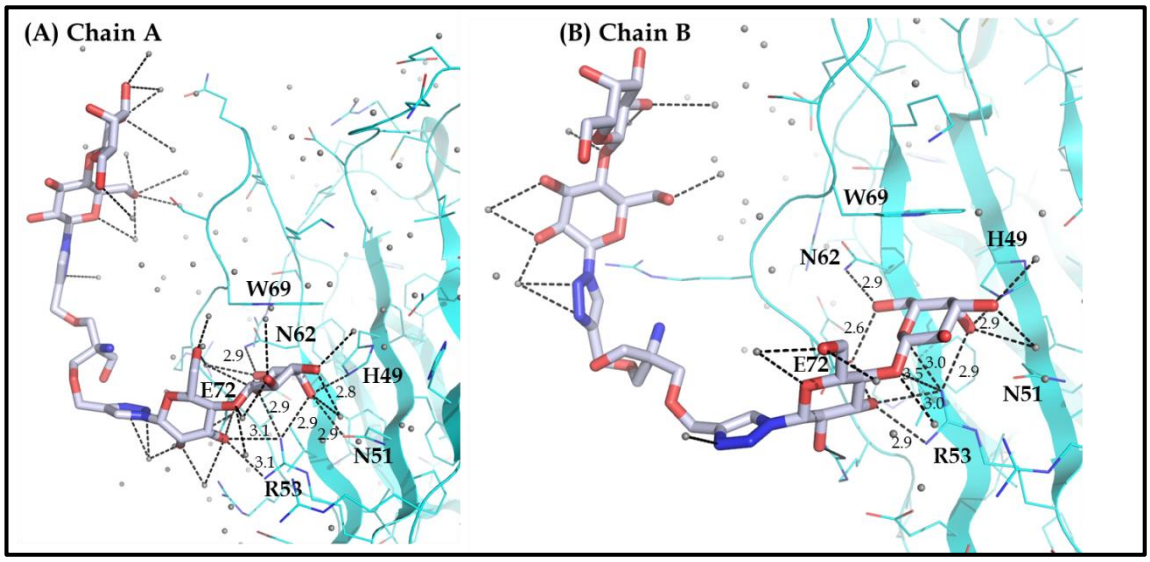


Figure 4. 19: CRD of hGal-7 in complex with D2-1.

Hydrogen bond interactions with both the chains of the dimer are shown. D2-1 is shown as sticks in pale blue and water molecules shown are as grey spheres. Distances are given in Å.

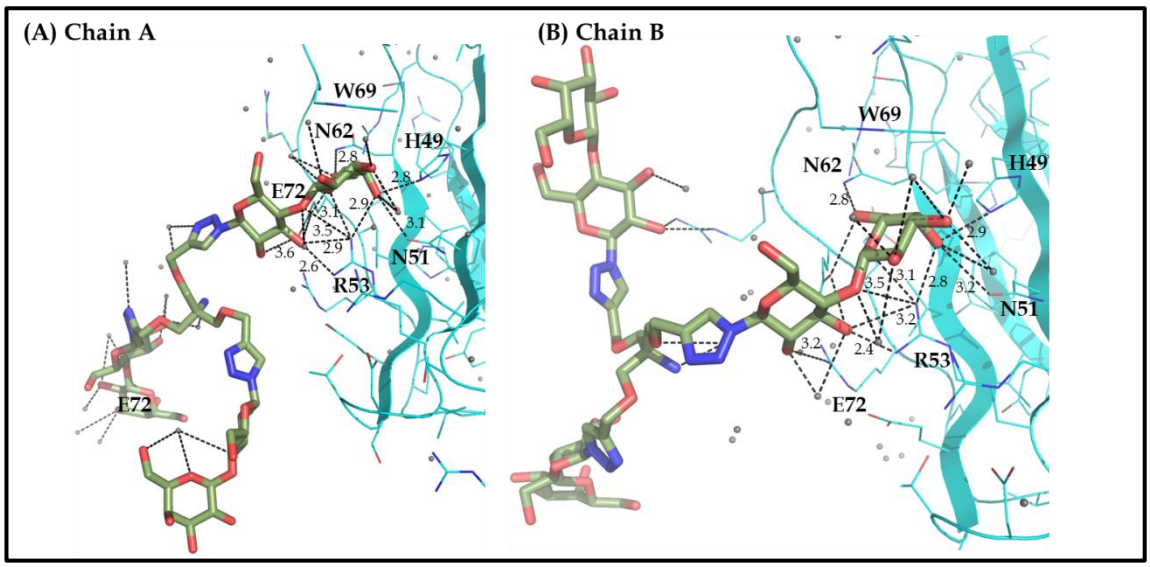


Figure 4. 20: CRD of hGal-7 in complex with D2-2.

Hydrogen bond interactions with both the chains of the dimer are shown. D2-2 is shown as sticks in smudge and water molecules are shown as grey spheres. Distances are given in Å.

The terminal galactosyl ring of D3 forms Π-Π stacking interactions with W69 and hydrogen bonds with other amino acids of the CRD. In addition to the hydrogen bonds [Table 4.8] observed between the terminal galactosyl ring of the dendrons and the CRD residues (H49, R53, N62, R71), a hydrogen bond between OZ3 and OGD of the diamido phenyl group forming the core of D3, is also observed [Figure 4.21]. This interaction along with the other non-bonded contacts that have been listed in Table 4.9, contribute to strengthening the binding of D3 to hGal-7.

In addition to the main arm of the ligand described above, a second arm with weak electron density was observed to interact with hGal-7 away from the CRD. This pseudo secondary binding site was slightly different between the two interacting hGal-7 dimers. With one of the partners, the ligand arm forms hydrogen bonds with S30 and non-bonded contacts bonds with S30, R31, D55 and D126 of hGal-7, while the symmetry related molecule of the ligand forms non-bonded contacts with G67 and R74 of a second hGal-7 partner [Figure 4.22]. These interactions occur at a pseudo-secondary binding site. In both the molecules of the dimeric hGal-7, D3 interacts with the hydrophobic pockets and

thus enables a 'tweezer-like' clinging of the protein by the ligand [Figure 4.22]. Hydrogen bonds mediated through water molecules are observed between N51 and R31 of hGal-7 and the ligand [Figures 4.17E-1, 4.17E-2, 4.22, 4.23].

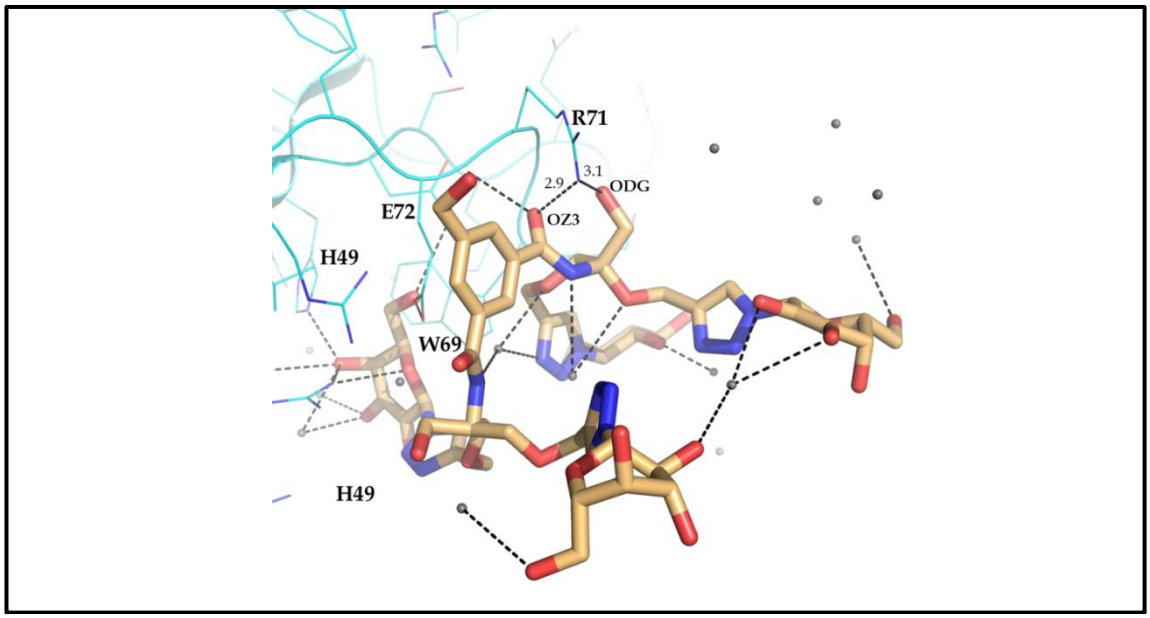


Figure 4. 21: CRD of hGal-7 in complex with D3.

OZ3 and ODG of the diamido phenyl group forming the core of the dendron D3 (shown as sticks in light orange), interacts with R71 of hGal-7. Water molecules are shown as grey spheres and distances are given in Å.

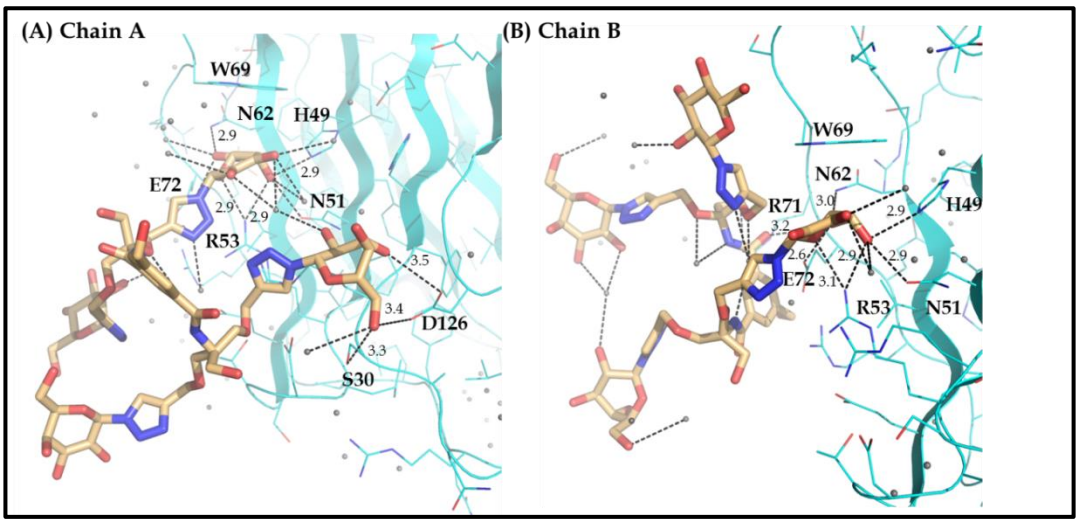


Figure 4. 22: CRD of hGal-7 in complex with D3.

H-bond interactions with both the chains of the dimer are shown. D3 is shown as sticks in light orange and water molecules are shown as grey spheres. Distances are given in Å.

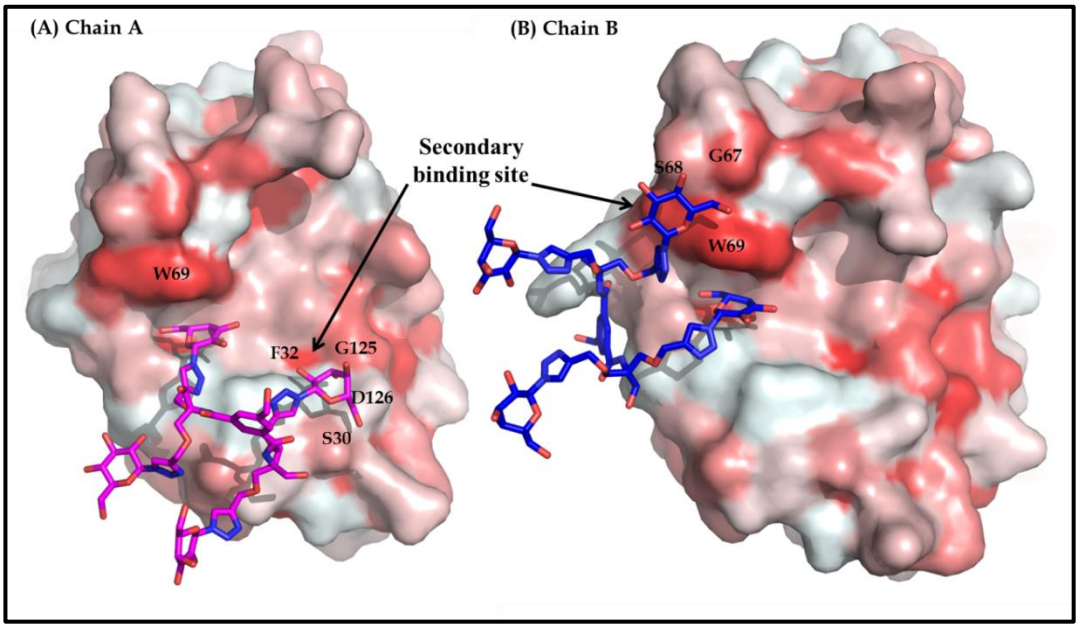


Figure 4. 23: Secondary binding site of D3 with hGal-7.

Two conformations of ligand binding are observed in the crystal structure of hGal-7 in complex with D3 (shown as sticks in pink and blue). The two arms of D3 bind to hydrophobic pockets of hGal-7 which results in 'tweezer-like' clinging of the protein by the ligand. hGal-7 is shown as surface and is coloured as per its hydrophobicity as described by Eisenberg *et al.* (1984) where red-through-white indicates hydrophobicity-through-hydrophilicity.

4.3.2.4. Cross-linking by the dendron

The recognition of dendrimeric compounds by hGal-7 results in cross-linking with the protein in each of the complexes described here. Complexes of hGal-7 with D1, D2-1 and D3 result in the cross-linking of two molecules of hGal-7. Interestingly the crystal form of D2-2 in complex with hGal-7 reveals that the dendron links three protein molecules [Figures 4.24 - 4.27]. The dimer partner of these hGal-7 molecules binds to another dendron thus, the dendrimeric units in each of these supramolecular assemblies act as a bridge between hGal-7 molecules.

The similarities in the electron density observed for the ligands D1 and D2-1 are also reflected in the similarity in the crystal packing of hGal-7 in complex with D1 and D2-1 [Figure 4.28, 4.29]. However, the crystal packing of the complex structures show some distinct differences [Figure 4.28 - 4.31]. It is observed that the distance between W69 at the CRD of adjacent hGal-7 molecules connected by the dendron increases with an increase in the dendron size. In the crystal structures of the dendron complexes, where electron density for the dendron connecting two molecules of hGal-7 (D1, D2-1 and D3) was observed, the distance between W69 at the CRD of adjacent hGal-7 is 16.3, 9.6 and 20 Å, respectively. However, in the hGal-7 - D2-2 structure, an increase in this distance to an average of 27.4 Å is observed. Thus the crystal packing in this crystal form provides sufficient volume for the ligand to bind to three different molecules of hGal-7.

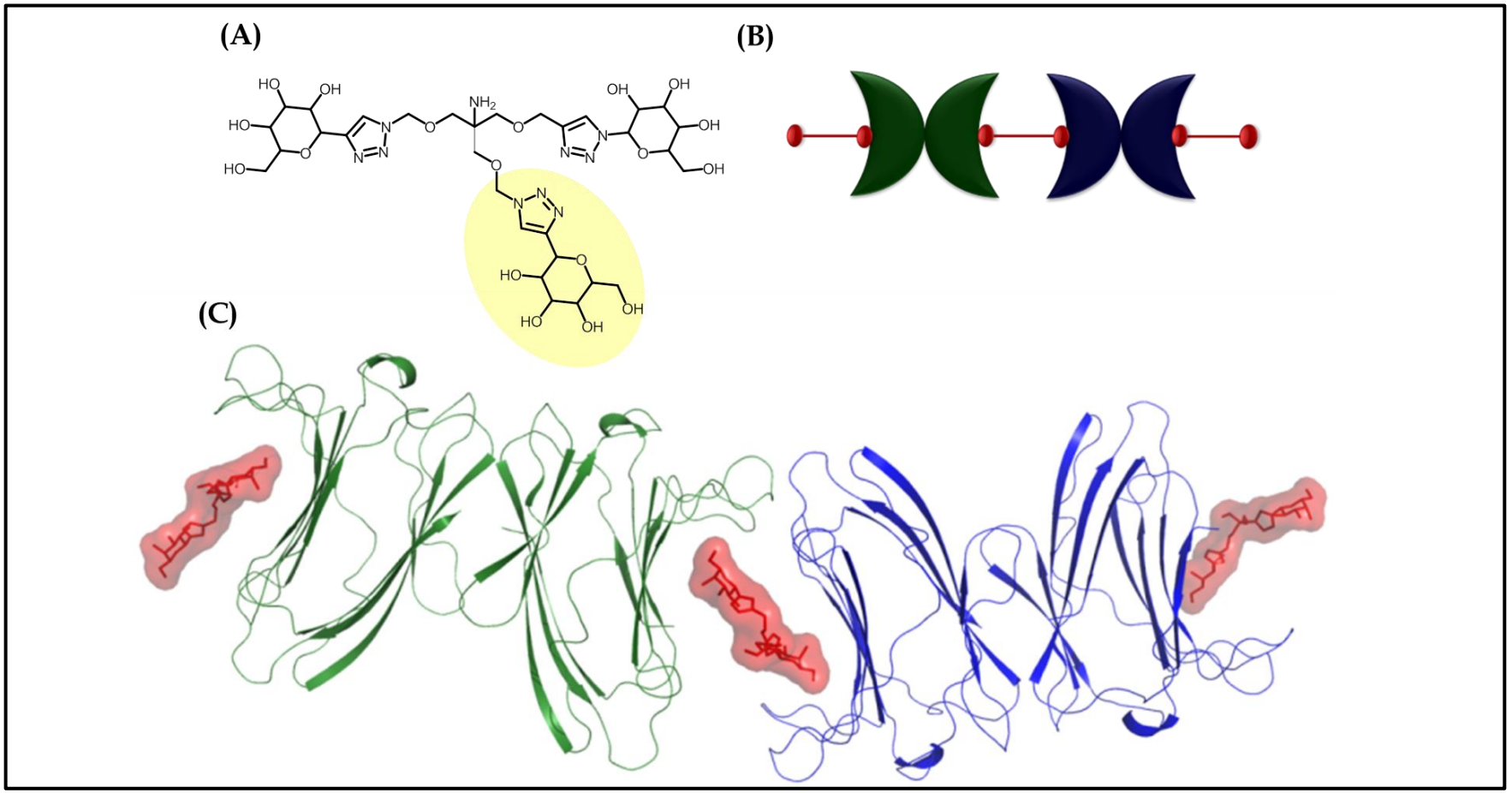


Figure 4. 24: Polymerisation forms of hGal-7 by D1.

(A) Chemical structure of D1 (The flexible arm of the ligand for which electron density is not observed is coloured in yellow). (B) Cartoon representation of polymerisation oh hGal-7 by D1. (C) Ribbon diagram of the hGal-7 dimer with D1 shown as sticks and surface.

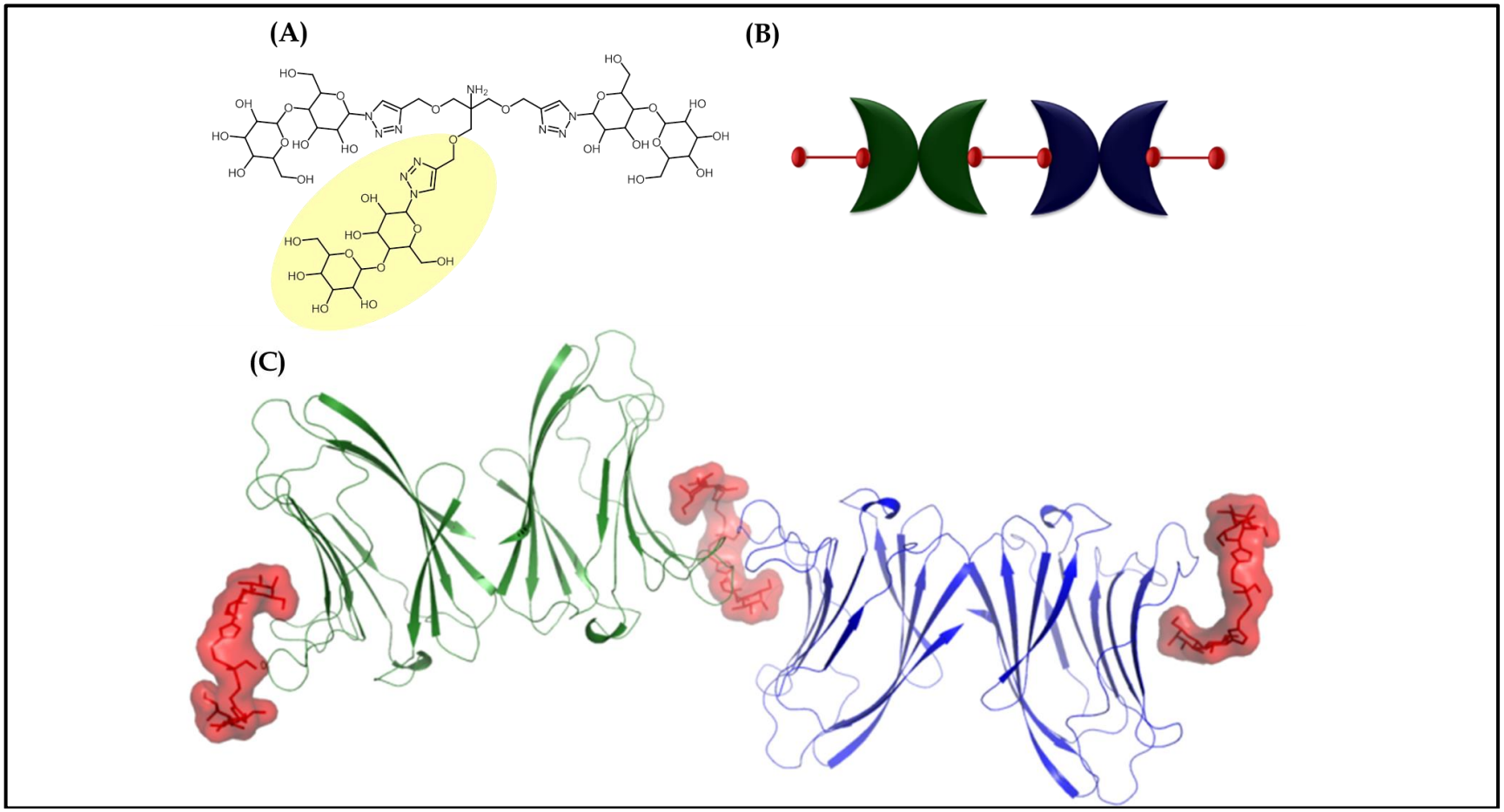


Figure 4. 25: Polymerisation forms of hGal-7 by D2-1.

(A) Chemical structure of D2-1 (The flexible arm of the ligand for which electron density is not observed is coloured in yellow). (B) Cartoon representation of polymerisation of hGal-7 by D2-1. (C) Ribbon diagram of the hGal-7 dimer with D2-1 shown as sticks and surface.

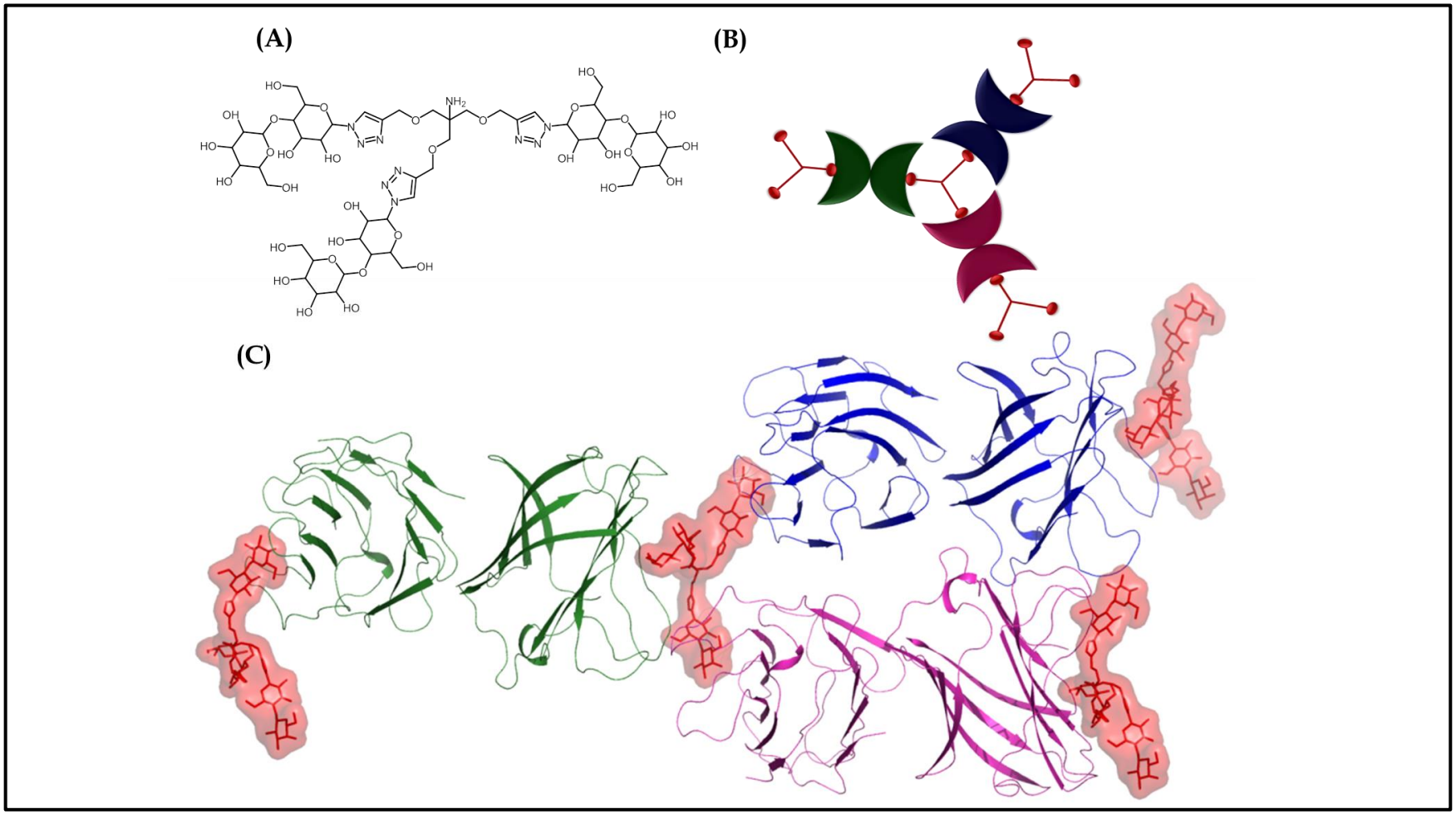


Figure 4. 26: Polymerisation forms of hGal-7 by D2-2.

(A) Chemical structure of D2. (B) Cartoon representation of polymerisation of hGal-7 by D2-2. (C) Ribbon diagram of the hGal-7 dimer with D2-2 shown as sticks and surface.

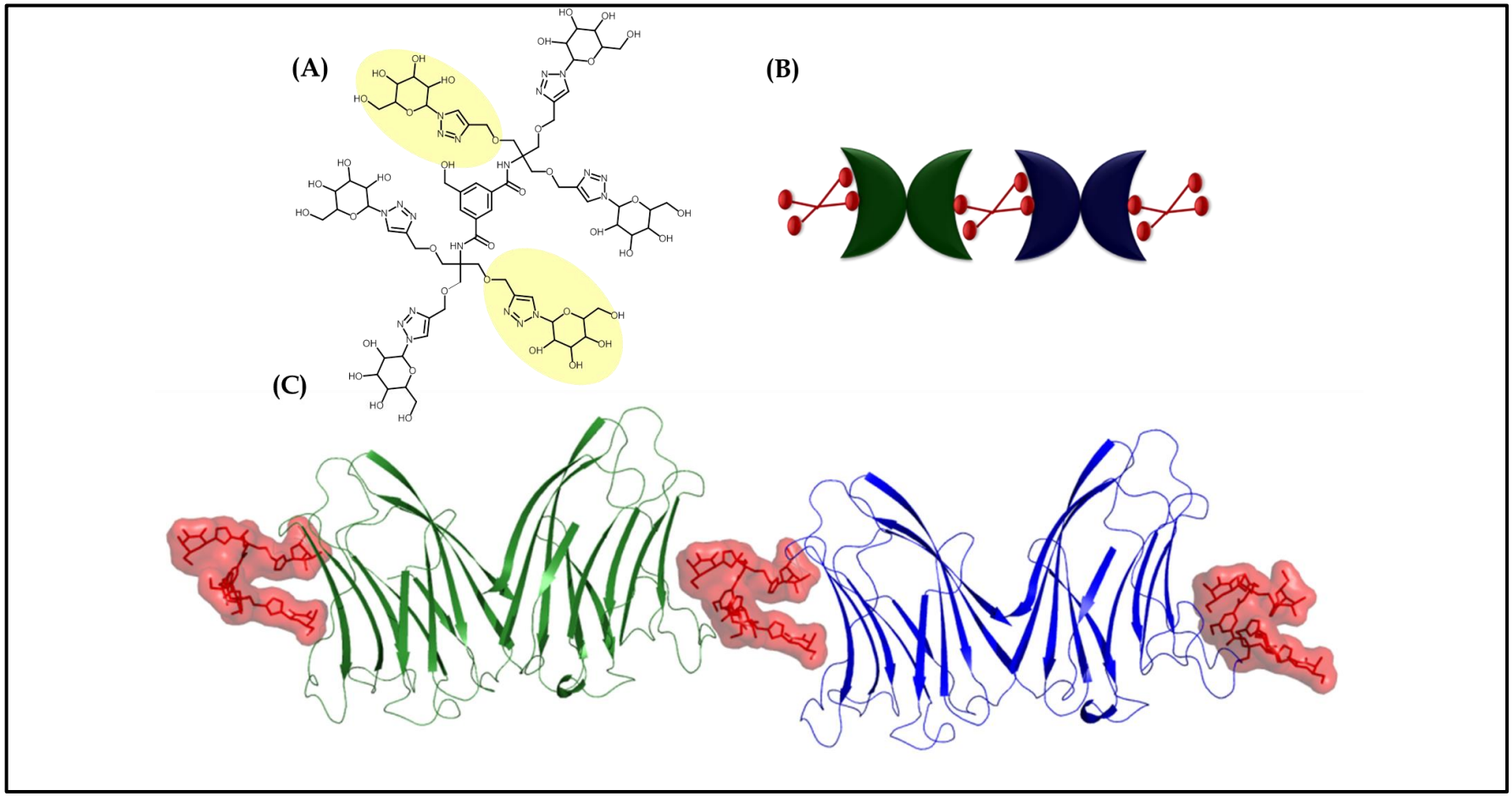


Figure 4. 27: Polymerisation forms of hGal-7 by D3.

(A) Chemical structure of D3 (The flexible arms of the ligand for which electron density is not observed is coloured in yellow). (B) Cartoon representation of the polymerisation of hGal-7 by D3. (C) Ribbon diagram of the hGal-7 dimer with D3 shown as sticks and surface.

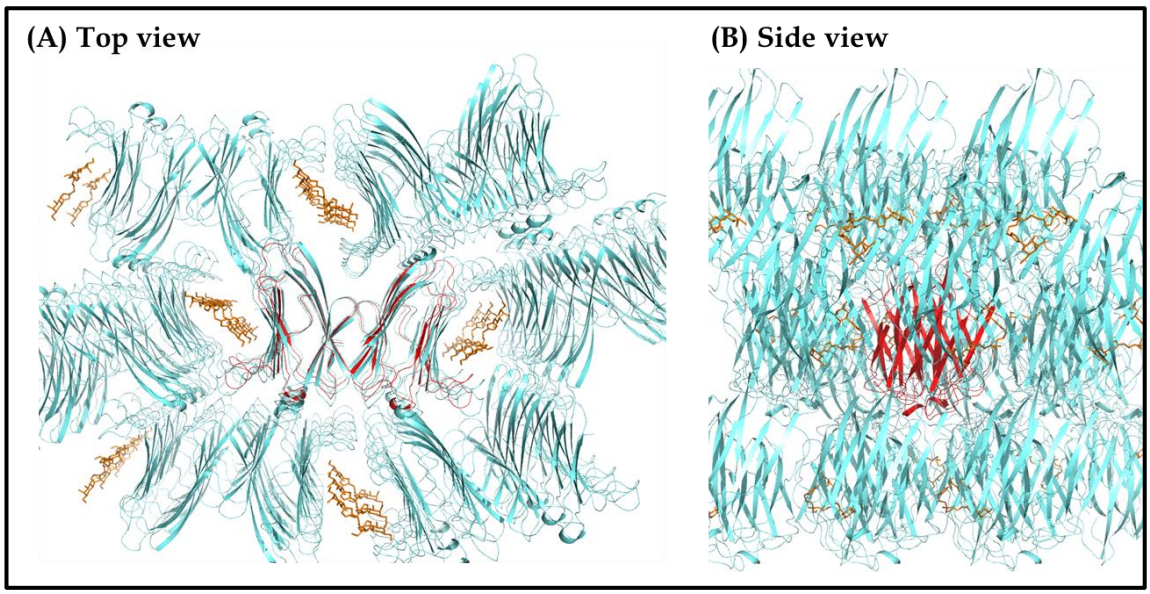


Figure 4. 28: Crystal packing of hGal-7 in complex with D1.

Molecules of the hGal-7 dimer are shown in red. Symmetry related molecules are shown in cyan and D1 is shown as sticks in orange.

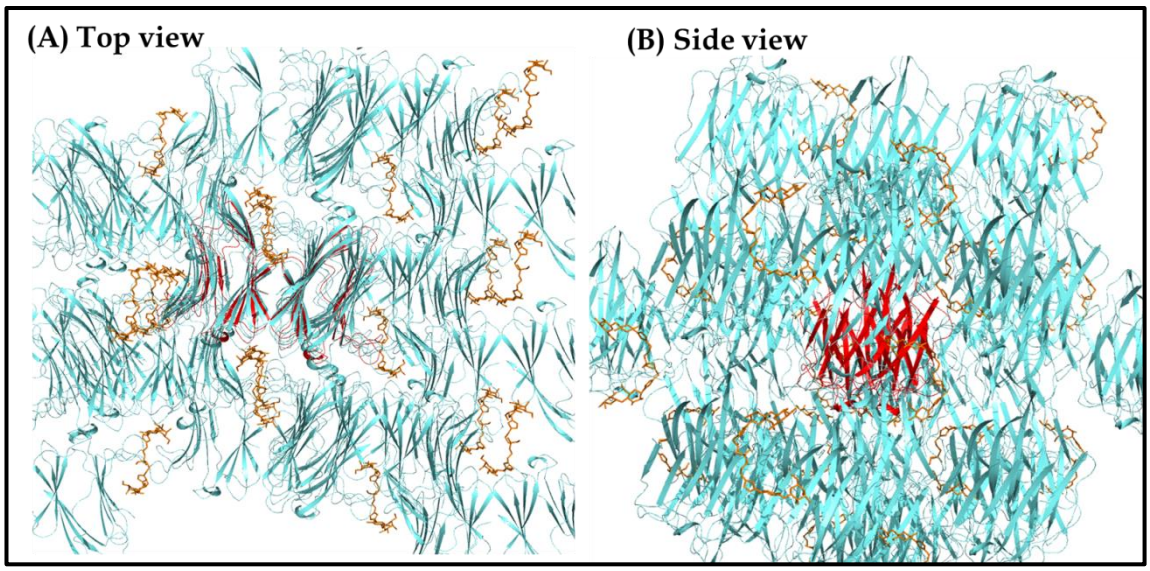


Figure 4. 29: Crystal packing of hGal-7 in complex with D2-1.

Molecules of the hGal-7 dimer are shown in red. Symmetry related molecules are shown in cyan and D2-1 is shown as sticks in orange.

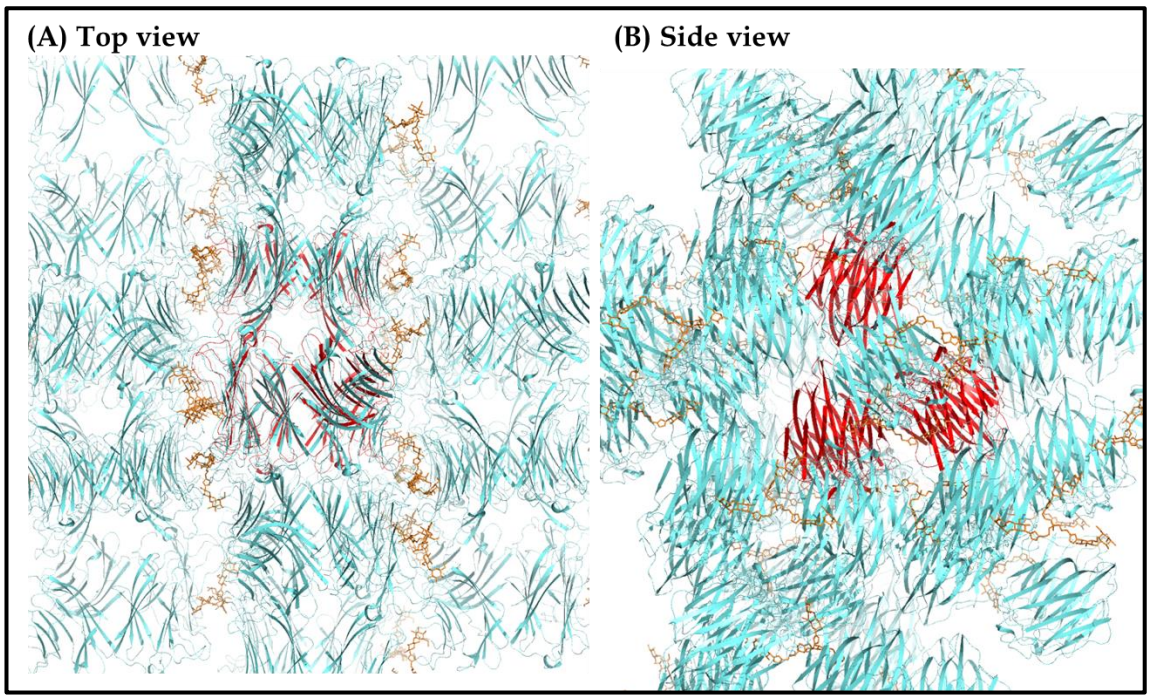


Figure 4. 30: Crystal packing of hGal-7 in complex with D2-2.

Molecules of the hGal-7 dimer are shown in red. Symmetry related molecules are shown in cyan and D2-2 is shown as sticks in orange.

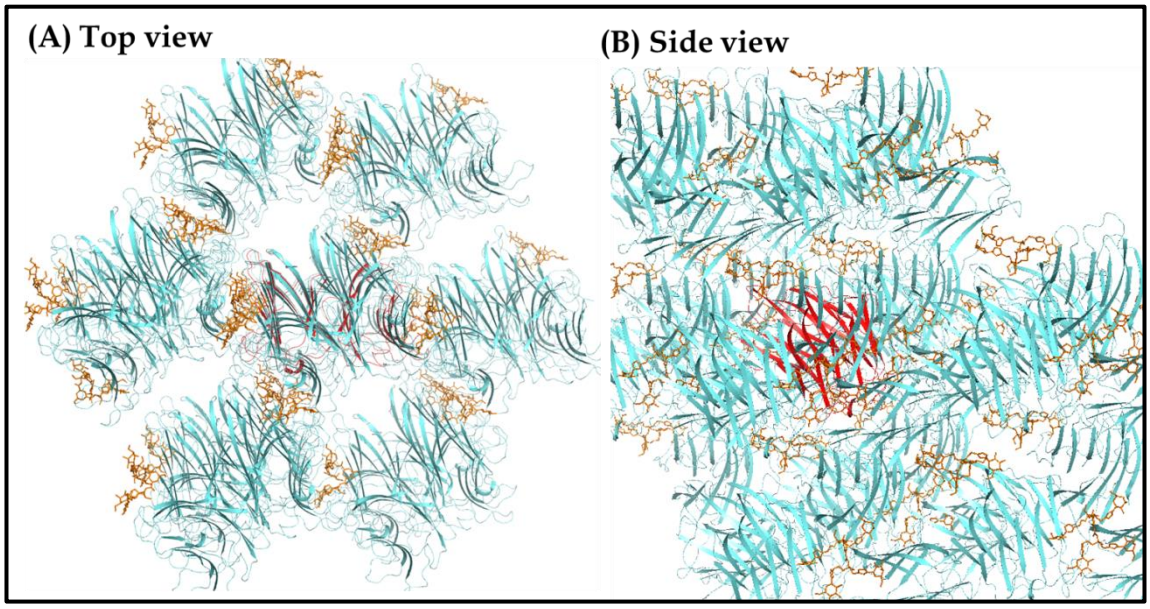


Figure 4. 31: Crystal packing of hGal-7 in complex with D3.

Molecules of the hGal-7 dimer are shown in red. Symmetry related molecules are shown in cyan and D3 is shown as sticks in orange.

4.3.4. Isothermal titration calorimetry studies

ITC studies were carried out to investigate the binding affinity of hGal-7 with the different dendrons used for crystallographic studies. A ligand concentration of 20 mM and a protein concentration of 5 mg/mL (0.33 mM) were determined as the optimum concentrations for the ITC experiments. Preliminary isothermal calorimetry data for D2 (K_d = 0.26 mM) interacting with hGal-7 gave an affinity constant several fold higher than that of GalNAc (K_d = 1.04 mM) and GAL (K_d = 2.17 mM) while the affinity for D1 (K_d = 3.12 mM) is comparable to GalNAc GAL (K_d = 1.04 mM) [Table 4.10 and figures 4.32, 4.33].

Table 4. 10: ITC data for GalNAc, GAL, D1 and D2 binding to hGal-7.
All experiments were performed at 15 ^{o}C. Values ± Stand deviation given.

	K_a (mM^{-1})	$-\Delta H$ (kJ mol^{-1})	$-\Delta G$ (kJ mol^{-1})	$T\Delta S$ (kJ mol^{-1})	n (sites/ monomer)	K_d (mM)	ΔS (J/ $mol^{-1}K$)
GalNAc	0.963 ± 0.05	4.3655 ± 0.73	16.55 ± 0.02	12.18 ± 0.72	4.241 ± 0.91	1.040 ± 0.05	42.28 ± 2.48
GAL	0.465 ± 0.06	1342.5 ± 146.4	59.66 ± 4.56	-1282.84 ± 141.81	0.005 ± 0.001	2.169 ± 0.28	-4452 ± 492.2
D1	0.321 ± 0.03	41.93 ± 2.36	13.83 ± 0.2	-28.10 ± 2.57	1.427 ± 0.03	3.119 ± 0.27	-97.51 ± 8.89
D2	3.926 ± 0.41	1847 ± 214.9	19.41 ± 0.41	-1827.59 ± 214.55	0.011 ± 0.002	0.256 ± 0.03	-6342.5 ± 744.5

β-Mercaptoethanol was used as a reducing agent in the buffer to ensure the presence of monomeric hGal-7 and to prevent its continuous cross-linking. However, the binding stoichiometry is ambiguous in this data [Table 4.10]. This could be due to the varying ratios of simple cross-linking of 'n' number of hGal-7 molecules (where n = number of arms of the ligand) and continuous cross-linking resulting in the formation of supramolecular assemblies.

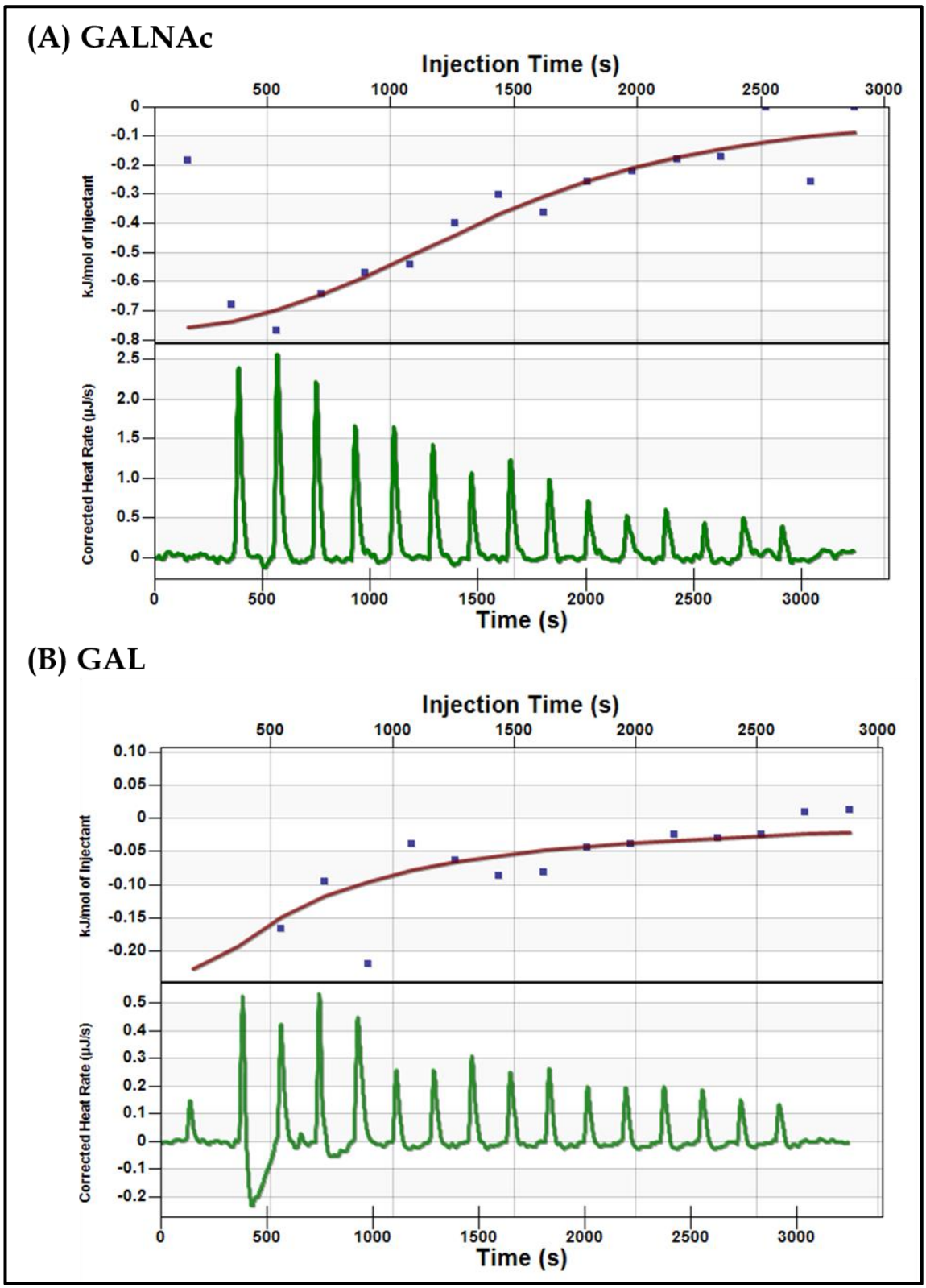


Figure 4. 32: ITC data for GALNAc and GAL in complex with hGal-7.

Bottom: thermograms obtained by injections of the compound. Top: corresponding integrated titration curves.

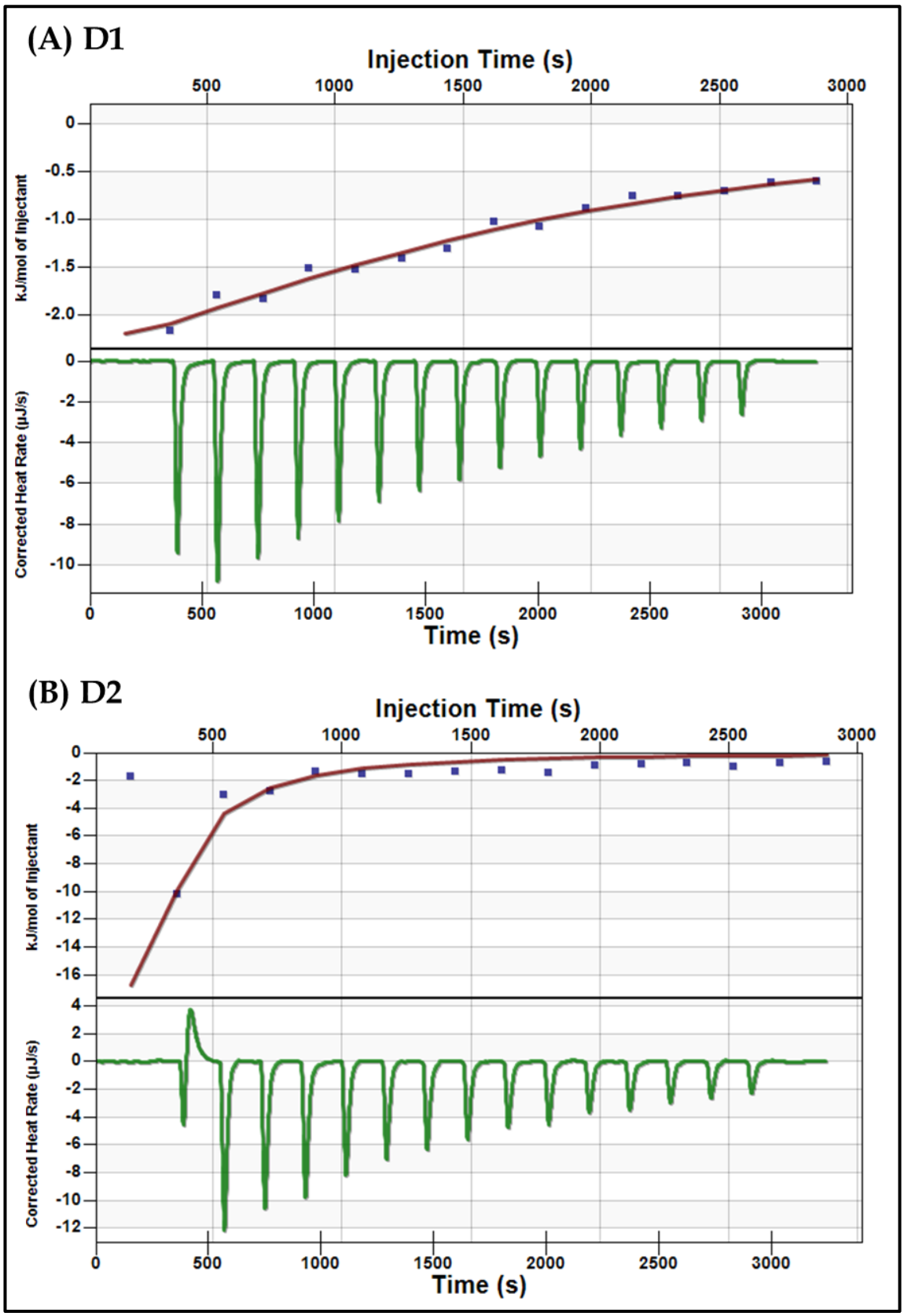


Figure 4. 33: ITC data for D1 and D2 in complex with hGal-7.

Bottom: thermograms obtained by injections of the compound. Top: corresponding integrated titration curves.

4.3.5. Dynamic light scattering experiments

DLS measurements were used to observe cross-linking of hGal-7 in solution by dendrons [Figure 4.34]. A single intensity peak was observed for native hGal-7. Upon ligand binding the intensity of this peak decreased from 100% for native hGal-7 to 12% for D1, 5.1% for D2, 46.9% for D3 and additional peaks for higher order structures (with diameters of D1: 1013 nm, 80% and 5560 nm, 7.4%; D2: 6733 nm, 95%; D3: 675 nm, 44.6% and 5292 nm, 8.4%) were observed [Table 4.11 and Figure 4.34]. The formation of these supramolecular assemblies in solution corroborate the crystallographic findings discussed in section 4.3.3.

In addition, a slight change in the size of the main hGal-7 peak in the presence of the ligands which appear to indicate changes in dimer size (Native: 7.82 nm; D1: 5.78 nm; D2: 8.99 nm; D3: 6.15 nm) upon ligand binding is observed [Table 4.11]. These fluctuations are reflected in the crystal structures of the complexes and are the result of changes in the dimeric interface induced by ligand binding at the CRD.

Table 4. 11: Representative dynamic light scattering measurements.

Sample	Peak	Size (diameter, nm)	Percentage (%)
Native hGal-7	1	7.82	100
hGal-7 – D1	1	5.79	12
	2	1013	80
	3	5560	7.4
hGal-7 – D2	1	8.99	5.1
	2	6733	95
hGal-7 – D3	1	6.15	49.9
	2	675	44.6
	3	5292	8.4

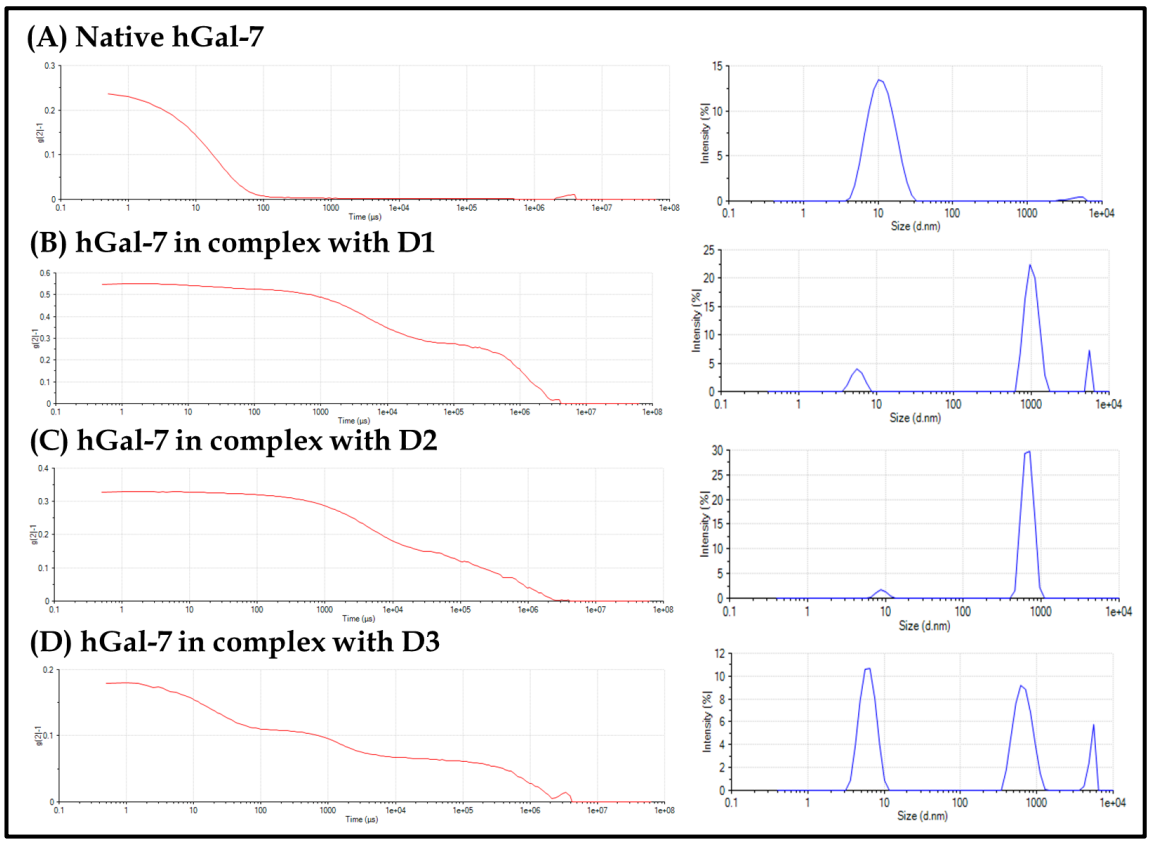


Figure 4. 34: Dynamic light scattering measurements.

Representative DLS data for native hGal-7 and hGal-7 in complex with dendrons D1, D2 and D3. Left panel: The autocorrelation function showing the exponential decay of the scattering signal. Right panel: The intensity graph gives the scattering intensity of the particles as a function of the size.

4.4. Discussion

Galectin–7, an important member of the β-galactoside-binding animal lectin family, has been widely implicated in cancer. The structural data of native hGal-7 and its known ligands have aided the development of synthetic carbohydrate based ligands. These enable us to study their activity through structure activity based relations. This wealth of knowledge has led to the development of specially designed dendrimers with terminal groups that are specific for their recognition by the CRD of hGal-7. The binding of dendrimers to galectins results in their polymerisation and helps us understand the implications and structural features of the formation of native supramolecular assemblies seen in other lectins. Galectins with a single CRD (prototype galectins) are known to dimerise and in some cases oligomerise upon binding to polysaccharides, which results in cell-cell interactions and signal transduction.

The different morphology [Figures 4.11] of the crystals observed in the crystallisation experiments can be attributed to the different crystal packing and supramolecular formation of hGal-7 in complex with dendrons of varying size and flexibility. Ligand flexibility, which could have resulted in disordered crystal formation, may help explain why the higher order dendrons (*i.e.* D4, D5 and D6) did not yield diffraction quality crystals.

Crystallographic results show that ligand binding at the CRD has a long range effect at the dimerisation interface. The binding of D2 and D3 to hGal-7 results in a higher RMS deviation of chain B compared to that of the native hGal-7 dimer [Tables 4.4, 4.5 and Figure 4.15]. The hydrogen bonds between residues R20–G102 and K98-F135 of the two monomers at the interface of the dimer are conserved in all the complex structures as well as in the native structure. Several additional hydrogen bonds are formed at the dimeric interface as a result of ligand binding at the CRD [Table 4.6]. Overall, the number of interactions between the residues at the dimeric interface in the complex structures remains comparable to that of the native structure, and even present slightly decreased surface contact areas. Therefore, although variations in the

dimer interface were observed upon ligand binding, they do not seem to offer additional stability to the dimer, in contrast to the observations made by Ermakova *et al.* (2013). While an effect of the crystallographic packing may not be excluded, the disparities seen at the dimeric interface more likely result from the flexibility within the supramolecular assemblies provided by the dendrons' extended arms.

The dendrimeric compounds are recognised by residues H49, N51, R53, N62, W69, E72 and R74 of the CRD of hGal-7. The interaction of the terminal galactosyl rings of the dendrons with the CRD are similar to that observed in the crystal structure of hGal-7 in complex with galactose. In addition to the recognition of the terminal galactosyl rings, the triazole ring (in D1 and D3) and the penultimate galactosyl ring (in D2), provide additional binding sites and further stabilise the complexes with hGal-7.

The pseudo secondary binding site observed in hGal-7 in complex with D3 appears to be a unique feature [Figure 4.35] compared to other structures of galectins bound to complex carbohydrates. Bourne *et al.* (1994) for example presented the first oligosaccharide structures of bovine Gal-1 cross-linked by N-acetyllactosamine units where the interactions occurred solely at the CRD. The rigidity of the glycosidic linkages in multi saccharide structures may prevent secondary interactions from taking place. Long and flexible dendrimers may enhance binding through cumulative interactions by interacting at the other hydrophobic pockets as observed in the case of D3 [Figure 4.23]. Additionally, the specific interaction of D3 with moderately conserved S30, R31, S68 and non-conserved D55 could be exploited to obtain specificity towards hGal-7 or other galectin sub-types [Figure 4.36]. The pseudo secondary site appears to only play a minimal role in the assembly of the supramolecules. It however indicates that better selectivity between lectins may be achieved by designing ligands able to target both the CRD and additional exosites.

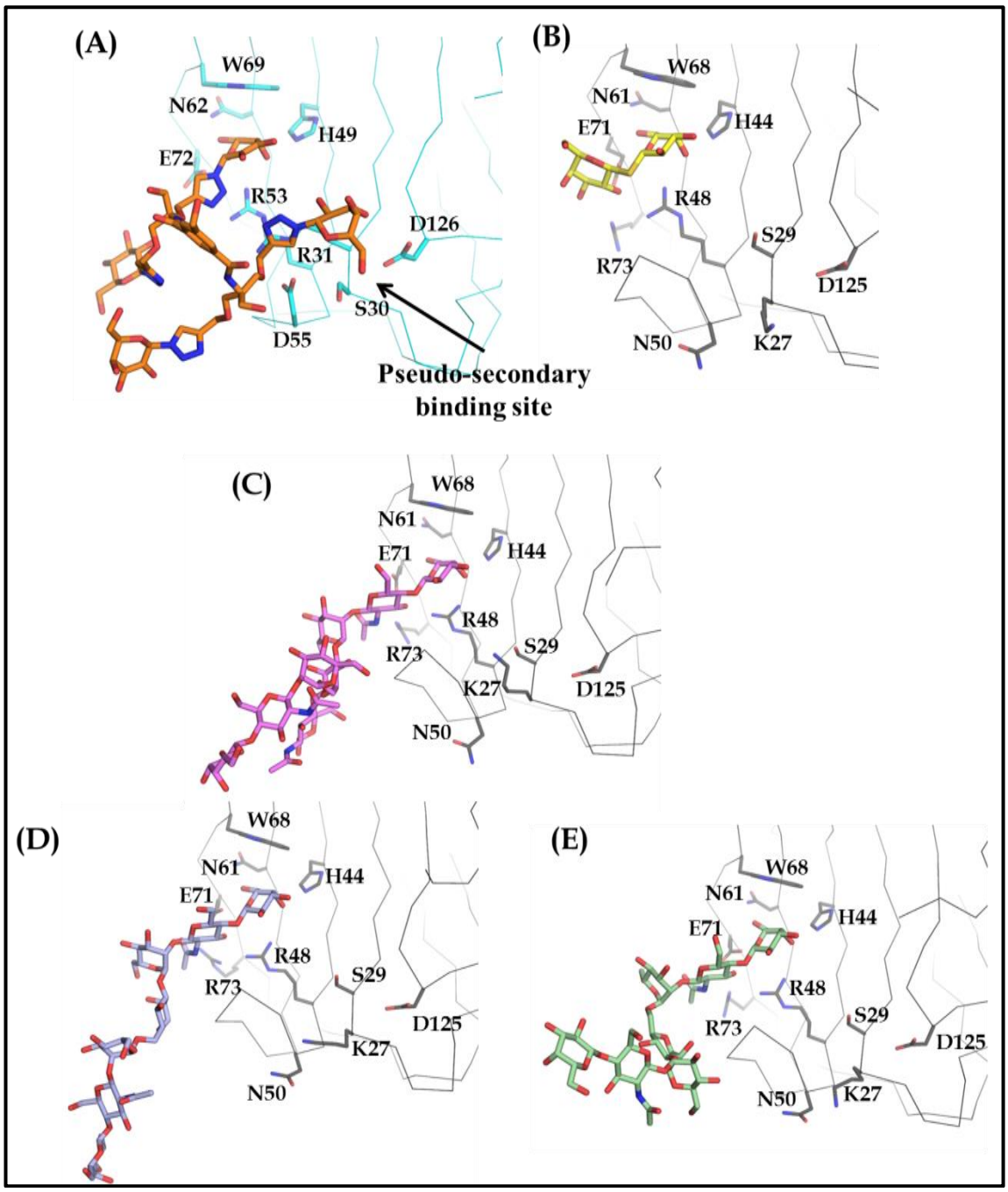


Figure 4. 35: Residues of the galectin CRD involved in cross-linking.

The pseudo- secondary binding site observed in the hGal-7-D3 complex is shown (hGal-7 in cyan and Gal-1 in grey**).** (A) hGAl-7 in complex with D3 (orange) (B) hGal-1 in complex with thiodigalactoside (yellow) (PDB code: 3OYW) (C-E) Bovine Gal-1 in complex with biantennary complex type saccharides (C) PDB code: 1SLA (violet) (D) PDB code: 1SLB (pale blue) (E) PDB code: 1SLC (pale green). Residues involved in the recognition and binding of the ligands are shown as sticks.

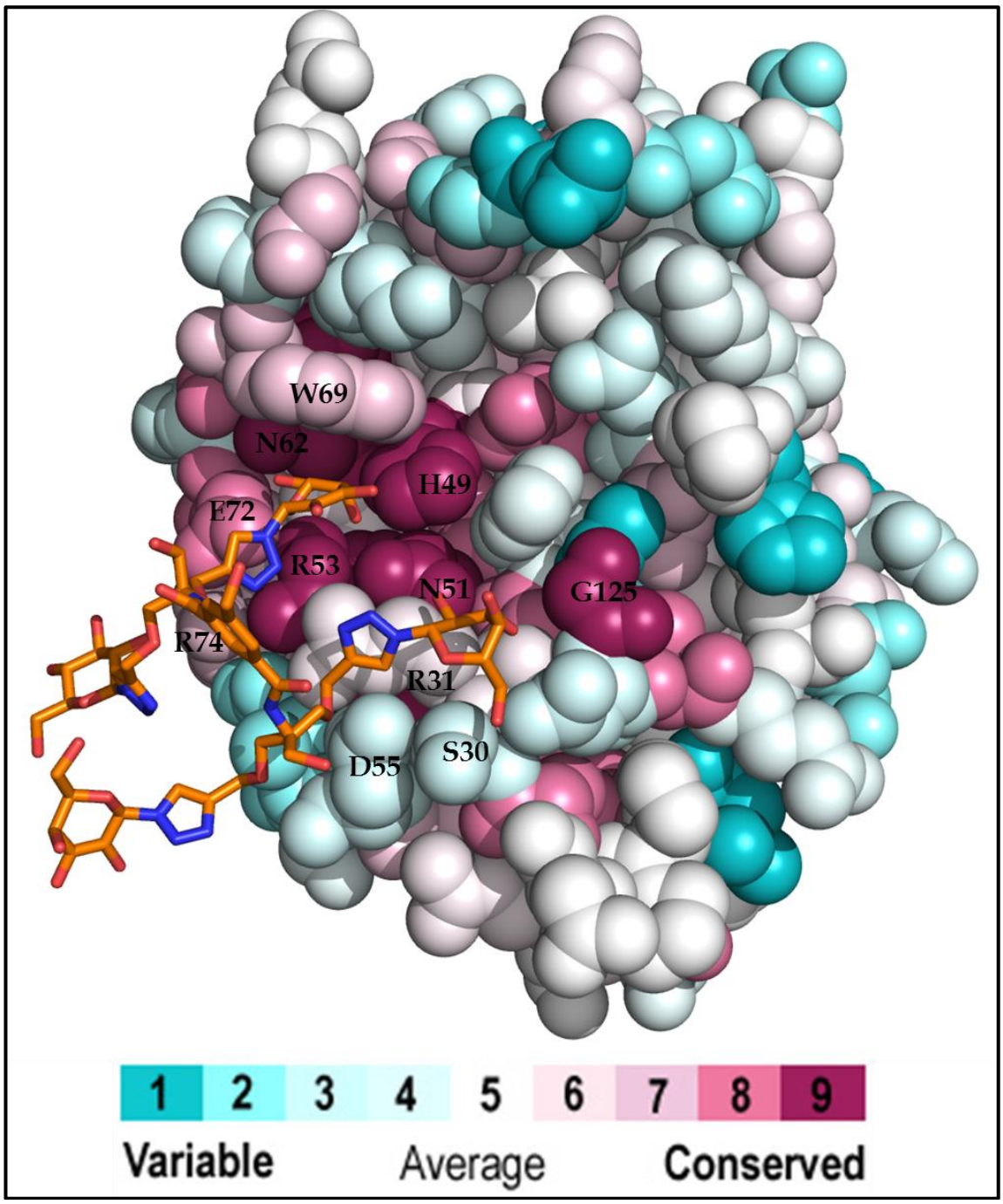


Figure 4. 36: ConSurf analysis of hGal-7 in complex with D3.

D3 is shown as stick in orange bound at the CRD. The amino-acids are coloured by their conservation grades using the colour-coding bar, with turquoise-through-maroon indicating variable-through-conserved.

The binding of multivalent carbohydrate based dendrons to hGal-7 promotes 'cross-linking' and results in the formation of large hetero-oligomers of hGal-7 and the multivalent dendrons. Depending on the size, flexibility of the dendrimeric ligand and the crystal packing of the complex, different modes of binding are observed. These results corroborate the glycoprotein clustering of hGal-3 by multivalent glycodendrimers studied using mass spectrometry and dynamic light scattering (Goodman *et al.*, 2014). The binding of the terminal dissacharide of D2 is analogous to the interaction of the disaccharide in the structure of hGal-1 in complex with thiodigalactoside (Stannard *et al.*, 2010) [Figure 4.35]. This similarity in the mode of binding highlights the potential use of dendrons and larger dendrimers for studying other members of the lectin family. Thiodigalactoside was suggested to have potential use as an adjuvant to promote vaccine stimulated immune responses against tumours (Stannard *et al.*, 2010).

On analysis of the ITC data, it is observed that the binding affinities of hGal-7 to GAL (K_d = 2.17 mM) and GALNAc (K_d = 1.04 mM) are much higher than that reported for lactose (K_d = 4.54 mM) and N-acetyl lactosamine (K_d = 5.88 mM) (Brewer, 2004). On comparison with monosaccharides, an enhancement of binding affinity of glycodendrimers with lectins ascribed to a "*cluster glycoside effect*" has been reported (Dimick *et al.*, 1999, Mouline *et al.*, 2014, Varki, 1993). This enhancement is observed in the binding affinity studies of D2 (K_d = 0.26 mM) with hGal-7. However, the binding affinity of D1 (K_d = 3.12 mM) is comparable to that of monosaccharides lactose, GAL and GALNAc.

The binding affinity data in conjuncture with the crystallographic results suggest that multivalent dendrons with additional galactosyl molecules at the termini (as seen in D2) may help enhance ligand binding. Future ITC and crystallographic studies with the higher-order dendrimers, will help further the understanding of supramolecular assembly formation of galectins. It will also be interesting to see the effect of the pseudo secondary site observed in the crystallographic results of hGal-7 in complex with D3 on the binding affinity of this dendron. Additionally, analytical ultracentrifugation studies or mass

spectrometric analysis to determine the exact size and polydispersity of higher order polymers in the hGal-7–dendron mixture may help in determining the stoichiometry of these reaction mixtures.

Dynamic light scattering was employed as a qualitative method to indicate the formation of supramolecular assemblies in solution. This showed that dendron-induced cross-linking of hGal-7, results in the formation of multiple species of varying sizes in solution. This polydispersity observed using DLS measurements, corroborate our crystallographic results in which two forms of cross-linking of hGal-7 by D2 and involvement of only parts of the dendrons, D1 and D3, are observed. Similar polydispersity in cross-linking of hGal-3 by glycodendrimers has been reported (Goodman *et al.*, 2014). Further analysis using techniques such as analytical ultracentrifugation or size exclusion chromatography with multi-angle light scattering may be required to further characterise the state of higher-order structures in solution.

In conclusion, hGal-7 polymerisation can be regulated by nucleating polymers such as the multivalent glycodendrimers used here. The crystal structures of the hGal-7 in complex with the dendrimeric compounds discussed will provide new tools to investigate the recognition of carbohydrates by galectins. The ability of these compounds to polymerise galectins will provide insights that will improve the general understating of their role in the formation of supramolecular assemblies and more generally, the lattice formation by lectins in cell signalling and signal transduction. The possibility of designing and synthesising dendrimers with specific functions, and their ability to be used in conjunction with nano-devices and liposomes make them ideal tools for such explorative studies.

Chapter 5. STRUCTURAL STUDIES ON SIRTUINS

Sirtuins, also known as the Sir2 family of proteins (*s*ilent mating type *i*nformation *r*egulator 2) were first identified in yeast where Sir2 came out on top in a screen for modulators of yeast lifespan (Kaeberlein *et al.*, 1999). They are NAD+ dependent deacetylases, but may have other related activities and are highly conserved in all organisms. The knowledge of their involvement in biological processes is steadily growing along with their implication as potential therapeutic target in various conditions such as ageing and age-related diseases including cancer. Due to their role in ageing and ageing related diseases, sirtuins have garnered the attention of the general public and researchers alike. Their unique enzymatic function, ability to modulate physiological processes on their own and by interplay with other members of the sirtuin family, and their diverse substrate recognition properties make them a fascinating family of enzymes. Work towards uncovering how their molecular activities explain their physiological outcomes in sirtuin influenced pathways has generated a great wealth of knowledge. However, further research aimed at understanding their dynamic molecular role in order to help design molecules to modulate their function is required.

Sequence alignment of all seven human sirtuins, reveals a 2.5% identity which includes 21 identical residue positions. Differences in the primary sequences which would be translated into the secondary structure, will provide insight into their varied affinity to diverse substrates and their role in normal cellular physiology.

The study of SIRT1, SIRT2 and SIRT7 was carried out to aid in understanding the structural implications on the dynamic role of sirtuins. This chapter describes the work carried out towards the structural studies of sirtuins. The cloning, expression, purification and crystallisation experiments carried out for each sirtuin (SIRT2, SIRT7 and SIRT1) are discussed separately in sections 5.1-5.3. The discussion of the results from these experiments is presented in section 5.4.

5.1. Sirtuin 2

5.1.1. Introduction

The human Sir2 ortholog, SIRT2, is the only main cytoplasmic isoform (Michishita *et al.*, 2005, Vaziri *et al.*, 2001) but its presence in the nucleus has also been reported (Bae *et al.*, 2004, Inoue *et al.*, 2007) and it appears to contain a nuclear export motif [Figure 5.1.1] (Cen *et al.*, 2011, Wilson *et al.*, 2006). SIRT2 deacetylates various substrates including α-tubulin and the histone H4, thereby regulating functions in cell cycle progression, cellular motility and differentiation (Cen *et al.*, 2011, North *et al.*, 2003, Vaquero *et al.*, 2006). Five splice variants of SIRT2 have been reported, Isoforms 1-5 have 389, 352, 369, 271 and 319 residues, respectively (UniProt, 2014). Isoform 1 is the canonical sequence and is expressed in the heart, liver and skeletal muscle. Isoform 2 is dominant in the brain and lacks an N-terminal extension to the catalytic domain (Afshar and Murnane, 1999, Frye, 1999). SIRT2 expression in the brain is indicative of its central role in this tissue and its contribution to pathologies of neurodegenerative diseases such as Parkinson's and Huntington's disease (de Oliveira *et al.*, 2012).

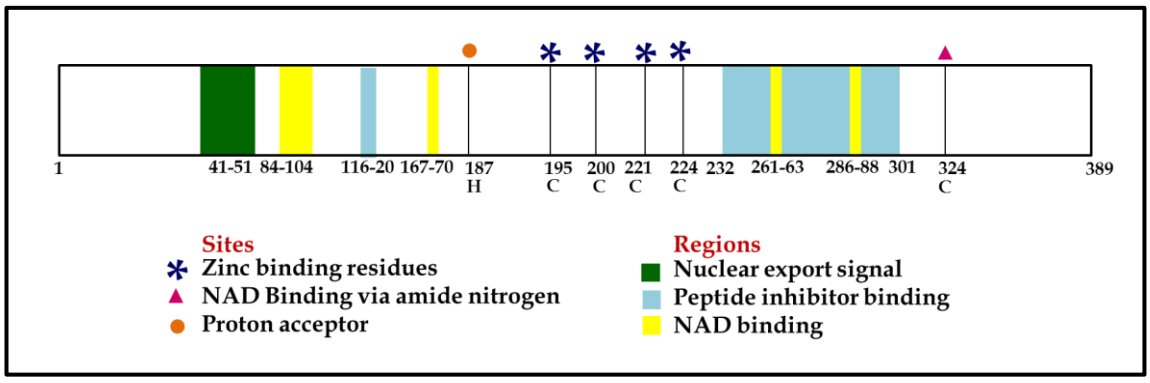


Figure 5.1. 1: Key features of the structure of SIRT2.

A schematic representation of the structure of SIRT2. The structurally important cysteine (C) residues involved in zinc binding are indicated along with the histidine (H) residue that behaves as a proton acceptor.

SIRT2 shows typical sirtuin architecture, in which a conserved catalytic core that is involved in substrate binding is flanked by N- and C-terminal extensions (discussed in section 2.5). SIRT3 is the closest homologue of SIRT2 while SIRT4

is the least similar [Table 5.1.1]. The N- and C-terminal regions appear to control sub-cellular localisation and play a role in catalytic activity (Flick and Luscher, 2012). SIRT2 activity can be modulated by C-terminal phosphorylation of residues S368 and S372 or by acetylation of lysines residues (Han *et al.*, 2008).

Table 5.1. 1: Overview of the results of a sequence comparison of SIRT2 with other human sirtuins.

ClustalO was used for sequence comparison (UniProt, 2014).

		Identity (%)	Identical positions
SIRT2	SIRT1	17.806	138
SIRT2	SIRT3	34.87	158
SIRT2	SIRT4	14.20	75
SIRT2	SIRT5	17.66	74
SIRT2	SIRT6	20.23	87
SIRT2	SIRT7	18.49	83

The first crystal structure of the catalytic core of SIRT2 revealed a typical oval shaped architecture as observed in other sirtuins composed of two globular subdomains, a Rossmann fold domain and a zinc-binding domain (Finnin *et al.*, 2001). The crystal structure of the catalytic core in complex with ADPR highlighted its differences from the apoform which included a 25^{o} rotation of the zinc-binding domain towards the Rossmann-fold domain which closed the co-substrate binding site around its ligand (Moniot *et al.*, 2013). The crystal structure of SIRT2 in complex with a macrocyclic peptide inhibitor S2iL5 showed the inhibitor induced open and closed domain movement of SIRT2 brought about by the movement of loop comprised of residues 289-304 (Yamagata *et al.*, 2014). The structural studies to date have mainly explored the catalytic core domain however little is known about the role and structure of the N- and C-terminal SIRT2 extensions.

Pharmacological inhibition of SIRT2 is an attractive approach for the treatment of neurological diseases and various cancers (discussed in section 2.5 and 2.6). Several inhibitors of sirtuins have been developed but have shown limited specificity to SIRT2 and have not yet been tested on other sirtuins and their

isoforms. The molecular details of ligand binding and the associated conformational changes are therefore of great interest (Moniot *et al.*, 2012).

The structural studies of SIRT2 presented here were aimed at understanding ligand recognition and binding to SIRT2. This would provide insights in SIRT2-specific and general sirtuin features that could be exploited for rational drug design.

5.1.2. Methods

5.1.2.1. Cloning, Expression and Purification

Table 5.1.2 summarises the constructs of SIRT2 that were designed for its expression with covalently linked affinity tags to help both expression and purification of the protein. The vector maps are provided in appendix A. The final optimised cloning, expression and protein purification protocols for each construct are given below.

Table 5.1. 2: Summary of SIRT2 constructs

Construct	Protein Insert (Residues)	Vector	Tag	Mol. Wt. (kDa)
SIRT2.1	34-356	pTriEx-4 Ek/LIC	N-term His and S-tag	41.5
SIRT2.4	34-356	pET22b	C-term His tag	37.6
SIRT2.6	34-389	pET-46 Ek/LIC	N-term His tag	41.5
SIRT2.7	34-356	pGEx-6P1	N-term GST tag	62.9

5.1.2.1.1. SIRT2.6

The gene coding SIRT2 (residues 34-389) was previously cloned into pET-46 Ek/LIC (Novagen, Darmstadt, Germany) with a cleavable N-terminal poly-histidine tag. This recombinant plasmid, named SIRT2.6, was then transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells. For protein expression, cells were grown in autoinduction (AI) media with ampicillin at 37 °C for 6 h and then incubated at 16 °C for 16 h. Cells were harvested after 22 h after inoculation by centrifugation (4 °C at 12,000 *g*, 10 min).

The expressed protein was purified using immobilised metal ion affinity chromatography (IMAC) followed by size exclusion chromatography (SEC). Cell pellets were resuspended in wash buffer (50 mM HEPES pH 7.3, 200 mM NaCl, 15 mM imidazole) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 75,000 g, 25 min). The supernatant was loaded on a 5 mL HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer. The

column was washed extensively and a gradient elution of 1-100 % over 50 mL (50 mM HEPES pH 7.3, 200 mM NaCl, 500 mM imidazole) was carried out. The eluted fraction was dialysed overnight in wash buffer at 4 °C and loaded on a gel filtration column (16/600 Superdex 200 pg; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM HEPES pH 7.3, 200 mM NaCl). Fractions were analysed by SDS-PAGE.

5.1.2.1.2. SIRT2.6a

The protein purified from the initial SIRT2.6 glycerol stocks was shown not to be the desired SIRT2 protein. The recombinant plasmid for SIRT2.6 was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequencing as described in section 3.4.2. Upon verification that the DNA sequence was correct, it was transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells and fresh glycerol stocks were prepared as described in Section 3.4.2. DNA was sequenced at Eurofins MWG Operon, Ebersberg, Germany using the T7 primer for forward sequencing and the TriExDOWN primer for reverse sequencing.

Expression trials were carried out by inoculating AI, LB and TB media with ampicillin and chloramphenicol (as described in Section 3.4.1) followed by induction with 1 mM IPTG. Samples were collected at regular intervals and protein expression was analysed using SDS-PAGE.

For large scale expression, cells were grown in TB media with ampicillin and chloramphenicol at 37 °C to an O.D. of 0.6–0.7. The cultures were then induced with 1 mm IPTG. Cells were harvested after 8 h of induction by centrifugation (4 °C at 12,000 g, 10 min). Induction of protein expression was analysed by SDS-PAGE.

SIRT2.6a was purified using IMAC followed by anion exchange (AIEX) chromatography. Cell pellets were resuspended in wash buffer 1 (50 mM HEPES pH 7.4, 200 mM NaCl, 15 mM imidazole) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 75,000 g, 25 min). The supernatant was loaded on a 5 mL

HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer 1. The column was washed extensively with wash buffer 1 before a wash with ~100 mL of 10 % elution buffer. The protein was eluted by gradient elution using 0-100% gradient over 50 mL with elution buffer 1 (50 mM HEPES pH 7.4, 200 mM NaCl, 500 mM imidazole) and the eluted fraction was dialysed overnight in wash buffer 2 (50 mM HEPES pH 7.4) at 4 °C. The dialysed sample SIRT2.6a (pI 5.84) was loaded on a manually packed 10 mL anion exchange column using quaternary amine (Q) sepharose beads (Sigma-Aldrich Co., Kent, U.K.) pre-equilibrated with wash buffer 2. The column was washed extensively and a gradient elution of 0-100 % over 100 mL using elution buffer 2 (50 mM Tris HCl pH 7.4, 500 mM NaCl) was carried out. The eluted fractions were analysed by SDS-PAGE and the desired fraction were dialysed overnight in final buffer (50 mM HEPES pH 7.3, 200 mM NaCl). Mass spectrometric analysis and western blot analysis, using anti-His antibody (Ab) and anti-SIRT2 Ab, were carried out as described in section 3.5.3. Antibodies were purchased from Sigma-Aldrich Co., Kent, U.K.

5.1.2.1.3. SIRT2.1

The gene coding SIRT2 (residues 34-356) was cloned into the pTriEx-4 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal poly-histidine and S-tag. This recombinant plasmid was named SIRT2.1.

Plasmid DNA was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequencing as described in section 3.4.2. Following DNA sequencing, the plasmid was transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells and fresh glycerol stocks were prepared as described in Section 3.3.2. DNA was sent for sequencing at Eurofins MWG Operon, Ebersberg, Germany using the T7 primer for forward sequencing and the TriExDOWN primer for reverse sequencing.

Expression trials were carried out by inoculating LB and TB media with ampicillin and chloramphenicol as described in Section 3.4.1 followed by

induction with 1 mM IPTG. Samples were collected at regular intervals and protein expression was analysed using SDS-PAGE.

For large scale expression, cells were grown in LB media with ampicillin and chloramphenicol at 37 °C to an O.D. of 0.6–0.7 and were then induced with 1 mM IPTG. Cells were harvested after 16 h of induction by centrifugation (4 °C at 12,000 g, 10 min). Induction of protein expression was analysed by SDS-PAGE.

SIRT2.1 was purified using IMAC followed by SEC. Cell pellets were resuspended in wash buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl, 10 mM imidazole) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 75,000 g, 25 min). The supernatant was loaded on a 5 mL HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer. The column was washed extensively and a gradient elution of 0-100 % over 50 mL (50 mM Tris HCl pH 7.4, 200 mM NaCl, 500 mM imidazole) was carried out. The eluted fraction was dialysed overnight in wash buffer at 4 °C and loaded on a gel filtration column (16/600 Superdex 200 Prep grade; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 Tris HCl pH 7.4, 200 mM NaCl). The eluted samples were analysed by SDS-PAGE. Tryptic digest of in-gel protein for mass spectrometric analysis and western blot analysis using anti-His Ab and anti-SIRT2 Ab were carried out as described in section 3.5.3. Antibodies were purchased from Sigma-Aldrich Co., Kent, U.K.

5.1.2.1.4. SIRT2.4

The gene coding SIRT2 (residues 34-356) was cloned into pET-22b (Novagen, Darmstadt, Germany) with a C-terminal poly-histidine. This recombinant plasmid was named SIRT2.4.

This plasmid was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequencing as described in section 3.4.2. Upon verification of the DNA sequence, the plasmid was transformed into *E. coli* BL21-CodonPlus

(DE3)-RIPL cells and glycerol stocks was prepared as described in section 3.3.2. DNA was sequenced at Eurofins MWG Operon, Ebersberg, Germany using the T7 primer for forward sequencing and the T7 term primer for reverse sequencing.

Expression trials were carried out by inoculating LB and TB media with ampicillin as described in Section 3.4.1 followed by induction with 1mM IPTG. Samples were collected at regular intervals and protein expression was analysed using SDS-PAGE.

For large scale expression, cells were grown in TB media with ampicillin at 37 ^{o}C to an O.D. of 0.6–0.7 and were then induced with 1 mM IPTG. Cells were harvested after 16 h of induction by centrifugation (4 ^{o}C at 12,000 g, 10min). Induction of the protein expression was analysed by SDS-PAGE.

Purification of SIRT2.4 was carried out using IMAC followed by SEC as described for SIRT2.1 in section 5.1.2.1.4. Tryptic digest of in-gel protein for mass spectrometric analysis and western blot analysis using anti-His Ab and anti-SIRT2 Ab were carried out as described in section 3.5.3. Antibodies were purchased from Sigma-Aldrich Co., Kent, U.K.

5.1.2.1.5. SIRT2.7

The gene coding SIRT2 (residues 34-356) was cloned into pGEx-6P-1 vector (GE Healthcare, U.K.) with a cleavable N-terminal GST-tag. This recombinant plasmid was named SIRT2.7.

The recombinant plasmid *SIRT2.6* was used for PCR amplification as described in section 3.4.2. Primers were designed such that a *BamH1* restriction site was incorporated into the DNA using the forward primer and an *EcoR1* restriction site was incorporated using the reverse primer. The primers used for PCR amplification are given in table 5.1.3, while the components of the reaction mixtures and the parameters for thermal cycling are given in tables 5.1.4 and 5.1.5, respectively.

Table 5.1. 3: Primers used for PCR amplification of SIRT2.7

Primer	Sequence (5′-3′)
SIRT2.7fwd	GAA TTC ATG GCA GCC GGG GGT CTG
SIRT2.7rev	CTC GAG TTA CGT CAC TTT CTT CCT

Table 5.1. 4: PCR reaction for SIRT2.7

Reaction	Tube 1 (µL)	Tube 2 (µL)	Tube 3 (µL)	Tube 4 (µL)
DNA template (SIRT2.6)(100ng/µL)	2	2	2	2
Primer (100 µM) *SIRT2.7fwd*)	1.5	1.5	1.5	1.5
Primer (100 µM) (*SIRT2.7rev)*	1.5	1.5	1.5	1.5
KOD polymerase (1U/µL)	1	1	1	1
10x Buffer	5	5	5	5
dNTPs (2mM each)	5	5	5	5
$MgSO_4$ (25 mM)	5	5	2	2
DMSO	2	-	2	-
Water (q.s. 50 µL)	27	29	30	32

Table 5.1. 5: PCR thermal cycling parameters for SIRT2.7

Cycle	Temperature (°C)	Time (min)
Initial denaturation	5	1
Denaturation	95	1
Annealing	57	1
Extension	72	2.5
Back to step 2 (Denaturation) for 35x		
Final extension	72	10
Hold at 4°C		

The PCR products were analysed by running the samples on an agarose gel as described in section 3.4.1. The final purified product was ligated with a pCR™-Blunt vector (Zero Blunt® PCR Cloning Kit, Invitrogen, CA, USA). The recombinant vector was transformed into *E. coli* TOP10 cells and the DNA was extracted as described in section 3.4.2. Double digestion experiments to test for the presence of an insert were performed by incubating the reaction mixture

[Table 5.1.6] for 1 h at 37 °C. FastDigest *BamHI* and *EcoRI* (Thermo scientific, MA, USA) were used for double digestion of the recombinant blunt vector (RBV) and pGEx-6P-1 vector (GE Healthcare, Uppsala, Sweden).

Table 5.1. 6: Double digestion set up for SIRT2.7.

	Test Double Digestion (µL)	Final Double Digestion (µL)	Vector Double Digestion (µL)
BamH1	1	4	4
EcoR1	1	4	4
RBV / (pGEx-6P-1 vector)	6	25	25
10x FastDigest Green buffer	1	5	5
Water	1 (q.s. 10 µL)	12 (q.s. 50 µL)	12 (q.s. 50 µL)

Once a dropout of the correct molecular weight was observed following test double digestions the blunt vector was sequenced using the M13(-21) primer. On verification of sequence, final double digestion of the RBV and pGEx-6P-1 vector was carried out [Table 5.1.6]. The Rapid DNA Ligation Kit (Thermo Scientific, MA, USA) was used for ligation of the insert and vector DNA. Ligation of the drop out (~1200 bp) from the RBV and the treated pGEx-6P-1 was carried out by incubating the ligation reaction (1 µL T4 DNA ligase, 1 µL treated pGEx-6P-1, 6 µL RBV drop out, 2 µL 5x rapid ligation buffer) for one hour at 21 °C. The recombinant plasmid was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequence analysis as described in Section 3.4.2. On verification of the sequence, the final recombinant plasmid was transformed into *E. coli* ArcticExpress (DE3) RIL cells (Agilent Technologies, CA, USA) and glycerol stocks were prepared as described in Section 3.4.2. Expression trials in LB media were carried out as per the manufacturer's instructions for *E. coli* ArcticExpress (DE3) RIL competent cells.

For large scale expressions, overnight cultures were grown in LB media with gentamycin and ampicillin. Large cultures were grown by inoculating modified LB media (+10 mM glucose) with overnight cultures containing no antibiotics.

These cultures were grown for 6 h at 30 ^{o}C and induced with 1 mM IPTG at 10 ^{o}C. Cells were harvested after 18 h of induction by centrifugation (4 °C at 12,000 g, 10 min). Induction of the protein expression was analysed by SDS-PAGE.

SIRT2.7 was purified using GST-tagged affinity chromatography (GST-AC) followed by GST-tag cleavage and SEC. Cell pellets were resuspended in PBS pH 7.3 (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) buffer and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 30,000 g, 30 min). The supernatant was loaded on a 5 mL GSTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash PBS pH 7.3 buffer. The column was washed extensively and a single step elution (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione) was carried out. Western blot analysis using anti-GST Ab and anti-SIRT2 Ab was performed to confirm the identity of the purified protein. Antibodies were purchased from Sigma-Aldrich Co., Kent, U.K. PreScission protease (GE Healthcare, Uppsala, Sweden) was added to the eluted protein (one unit protease cleaves 90% 100 µg of GST fusion protein) and dialysed overnight in 50 mM Tris HCl pH 8.0, 10 mM reduced glutathione at 4 ^{o}C. The GST cleaved protein was loaded on a gel filtration column (16/600 Superdex 200 pg; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl). The eluted samples were analysed by SDS-PAGE and by tryptic in-gel digest mass spectrometric analysis as described in section 3.5.3.3.

5.1.2.2. SIRT2 activity assay

The *Fluor-de-lys*®-SIRT2 fluorometric drug discovery assay kit (Enzo Life Sciences, NY, USA) was employed to measure the activity of purified SIRT2 protein. This assay was also used to determine the inhibitory capacity of sirtuin inhibitors used in the crystallisation experiments.

Principle of Fluor-de-lys®**-SIRT2 assay**

The assay involves the incubation of *Fluor-de-lys*® substrate with SIRT2 resulting in its NAD^+-dependent deacetylation. The deactylated *Fluor-de-lys*® substrate can then react with the *Fluor-de-lys*® developer II to produce a fluorophore [Figure 5.1.2]. The fluorophore is excited at 360 nm and the emitted light is detected on a fluorometric plate reader. NAD^+ is consumed in the reaction to produce nicotinamide (NAM) and O-acetyl-ADP-ribose (AADPR).

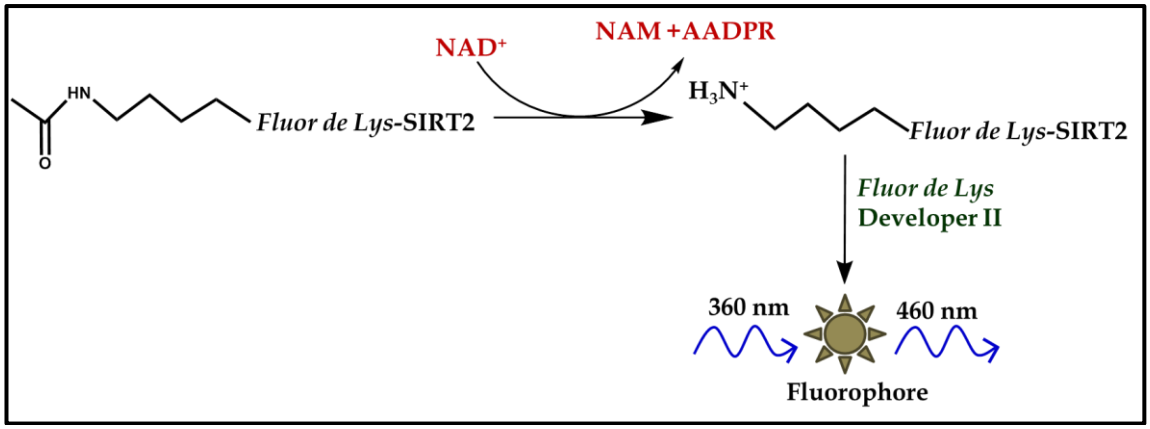


Figure 5.1. 2: Principle of SIRT2 fluorescent activity assay.

Experimental technique

The experiments were performed as per manufacturer's instructions in a 96-well NUNC microplates (Thermo Fisher Scientific, MA, USA). Human recombinant SIRT2 provided in the kit was used as a control. The FLUOstar Omega Multi-Mode Microplate Reader with a CCD-based Spectrometer (BMG LABTECH, Germany) was used for reading the plate. Excitation at a wavelength of 360 nm and detection of the emitted light at 460 nm was used for reading the plate.

SIRT2 activity assay

Test samples (SIRT2.1, SIRT2.4, SIRT2.6 and SIRT2.7) were dialysed in protein buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl) and initial stock solutions of 5 mg/mL, 2 mg/mL and 1 mg/mL were prepared. The control samples were

prepared at similar concentrations to the test samples. The substrate mix and developer mix were prepared as per manufacturer's instructions. The test samples/control and NAD^+-substrate mix [Table 5.1.7] were mixed and incubated at 37 ^oC. After 1 h, the developer and the nicotinamide stop solution (Developer mix) were added to the reaction mixture [Table 5.1.7]. The plate was read using the fluorescence spectrometer after 30-45 min of incubation.

Table 5.1. 7: SIRT2 activity assay

Sample	Assay buffer (µL)	Protein (µL)	PB (µL)	Substrate mix (µL)	Developer mix (µL)
Blank	5	-	5	10	20
Test sample	5	5	-	10	20
Control	5	5	-	10	20

SIRT2 inhibitor assay

The SIRT2 inhibition assay was performed similar to the SIRT2 activity assay. The inhibitors AGK2, cambinol (CAM), salermide (SEL) and suramin at initial concentrations of 200, 20, 2 and 0.2 mg/mL were prepared in assay buffer. All compounds were purchased from Calbiochem, U.K. and dissolved in DMSO (except suramin which was dissolved in water) to prepare 10 mM stock solutions. A DMSO control was assayed along with the inhibitors. SIRT2.6 at 0.5 mg/mL was incubated with the inhibitors for 1 h at 37 ^oC. The test sample (SIRT2.6 + inhibitor mix) and the NAD^+ - substrate mix [Table 5.1.8] were incubated at 37 ^oC. After 1 h, the developer and the nicotinamide stop solution (Developer mix) were added [Table 5.1.8] and incubated at 37 ^oC. The plate was read using a fluorescence spectrometer after 30-45 min of incubation.

Table 5.1. 8: SIRT2 inhibitor assay

Sample	Assay buffer (µL)	Protein (µL)	Inhibitor (µL)	Protein Buffer (µL)	Substrate mix (µL)	Developer mix (µL)
Assay Blank	5	-		5	10	20
Inhibitor Blank	5	5	-	-	10	20
Test sample	-	5	5	-	10	20

5.1.2.2. Crystallisation of SIRT2

Crystallisation trials for SIRT2 started with screening of sparse-matrix commercial conditions (Section 3.7.2). The vapour diffusion sitting drop method was the preferred crystallisation technique. A variety of conditions were tested at different temperatures (4 ^{o}C, 16 ^{o}C and 21 ^{o}C), with drops containing protein and reservoir solution in ratios of 1:1, 1:2 and 1:3.

The protein obtained from constructs SIRT2.1, 2.4, 2.6 and 2.7 were used for setting up crystallisation trials as described in section 3.7.2. Initial crystallisation trials of native SIRT 2 were set up in their respective final purification buffers. Protein concentrations of 2 mg/mL, 6 mg/mL, 8 mg/mL, 10 mg/mL and 12 mg/mL were used for these experiments.

In addition to the commercial screens available, optimisation screens based on the crystallisation conditions from the reported SIRT2 structures were also used (Finnin *et al.*, 2001, Moniot *et al.*, 2013). These included 96-well screens designed based on varying concentrations of PEG10000, ammonium acetate and NaCl. Additional conditions varying the ionic strength of the crystallisation conditions by altering the pH and concentration of bis-tris propane and bis-tris methane were employed.

A similar sparse matrix screening of SIRT2.6 and SIRT2.7 in complex with the substrates NAD^{+} and ADPR (Sigma-Aldrich Co., Kent, U.K.) as well as the inhibitors CAM, SEL, AGK2 and surmain were also set up. Protein

concentrations of 2 mg/mL, 6 mg/mL and 8 mg/mL were used and protein to ligand ratio was either 1:2 or 1:4.

In addition to the sitting drop method, modified microbatch or crystallisation under oil, were attempted to crystallise native SIRT2.6. The crystallisation robot Oryx8 (Douglas Instruments, U.K.) was used to set up 96-well plates using PACT *premier* HT-96, JCSG-*plus* HT-96, Morpheus HT-96, PGA Screen HT-96 and Heavy + Light twin pack HT-96 screens (Molecular Dimensions, U.K.).

5.1.3. Results

5.1.3.1. SIRT2 construct SIRT 2.6

SIRT2.6 is truncated at the N-terminus compared to native SIRT2 and comprises residues 34-389. The recombinant gene was cloned into the pET-46 Ek/LIC vector to enable protein expression with a covalently linked His-tag. The expressed protein is expected to have a theoretical mass of 41.5 kDa.

Initial expression in TB media followed by IMAC and SEC of SIRT2.6 resulted in two proteins being observed by SDS-PAGE at ~55 kDa and ~66 kDa [Figure 5.1.3 and 5.1.4]. Western blot analysis using anti-His Ab and anti-SIRT2 Ab failed to detect any protein, which confirmed neither the purified proteins were the desired SIRT2.

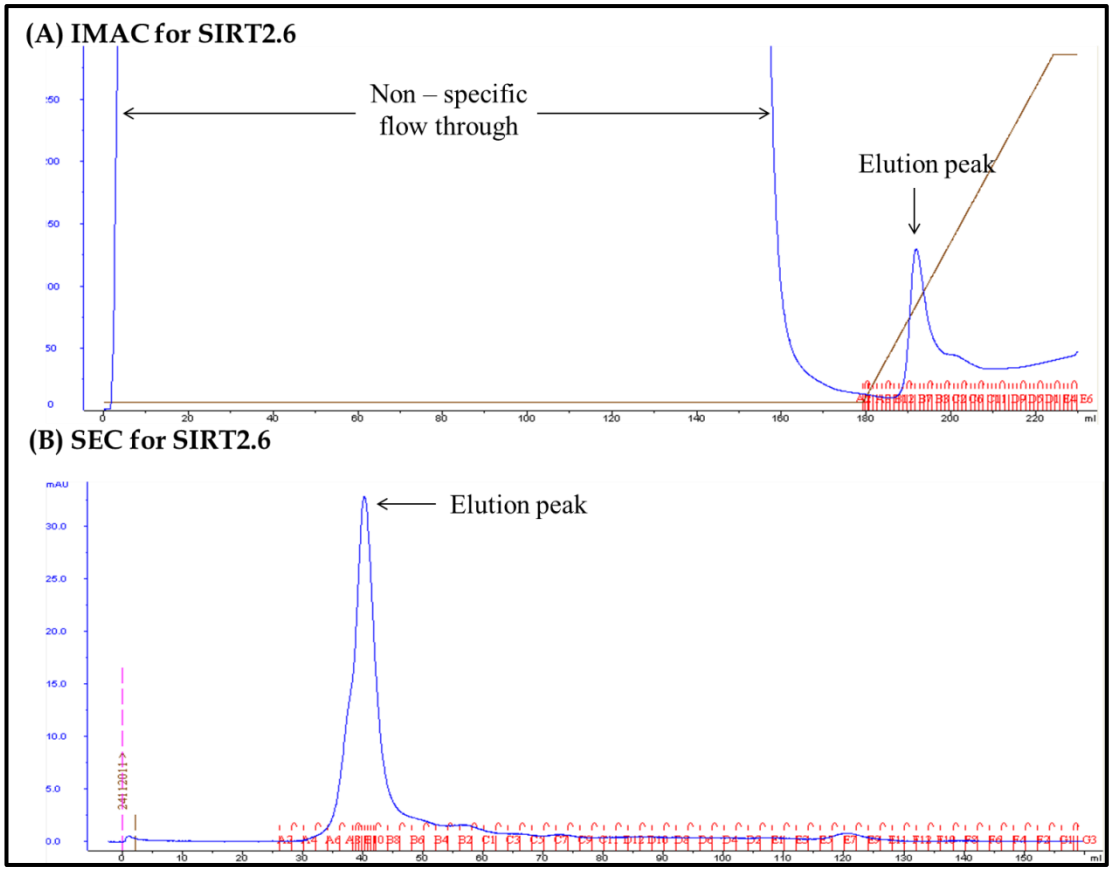


Figure 5.1. 3: SIRT2.6 purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT2.6 followed by (B) SEC using a superdex 200 column. The elution peaks are indicated in the figure. Blue line represents the absorbance at A_{280} and the brown line represents the percentage of elution buffer.

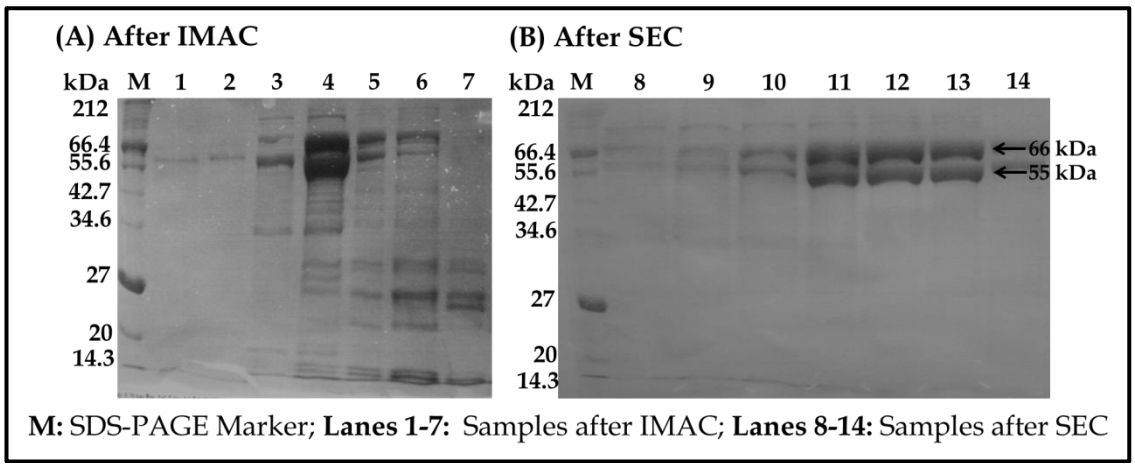


Figure 5.1. 4: Example of SIRT2.6 purification analysed by SDS-PAGE.

SIRT2.6 purification analysed by SDS-PAGE (A) samples from IMAC (lanes 1-7) followed by (B) SEC (lanes 8-14). The molecular weights of the SDS-PAGE marker (M) are labelled.

The verified plasmid isolated from the *E. coli* TOP10 cells glycerol stock was used for verification of DNA sequence (referred to as SIRT2.6a). Positive results for forward and reverse DNA sequencing were obtained. The alignment of the translated DNA sequences of SIRT2.6a with SIRT2 is given are in appendix E1.1 and E1.2. The recombinant plasmid was successfully transformed into Rosetta-gami™ 2(DE3)pLysS cells and expression trials were performed in LB, TB and AI media. Soluble expression was most favourable in TB and this media was selected for large scale expression of SIRT2.6a [Figure 5.1.5].

Initial IMAC purifications of SIRT2.6a resulted in a 30 kDa impurity being co-purified with the desired protein. Washing the HisTrap FF column after loading with ~100 mL of 10 % elution buffer enabled the removal of this impurity. Successful purification of SIRT2.6a was achieved using IMAC followed by AIEX [Figure 5.1.6 and 5.1.7]. A yield of 20-25 mg of SIRT2.6a per litre of cultures was obtained using this method. The identity of the protein was confirmed by intact protein mass spectrometry and western blot analysis using anti-His Ab and anti-SIRT2 Ab [Figure 5.1.8 and 5.1.9]. The SIRT2 activity assay and crystallisation screening for this protein is discussed in sections 5.1.3.4 and 5.1.3.5, respectively.

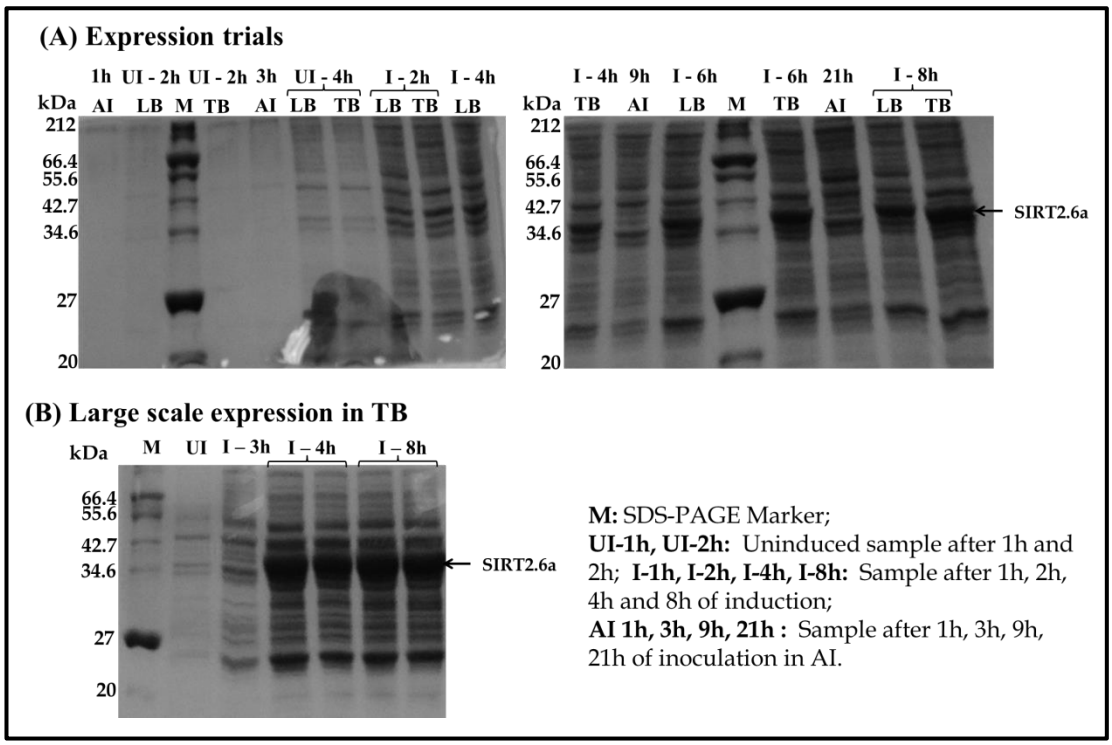


Figure 5.1. 5:Example of SIRT2.6a expression analysed by SDS-PAGE.

(A) Samples from SIRT2.6a expression trials (B) samples from large scale expression of SIRT2.6a in TB. The molecular weights of the SDS-PAGE marker (M) are labelled.

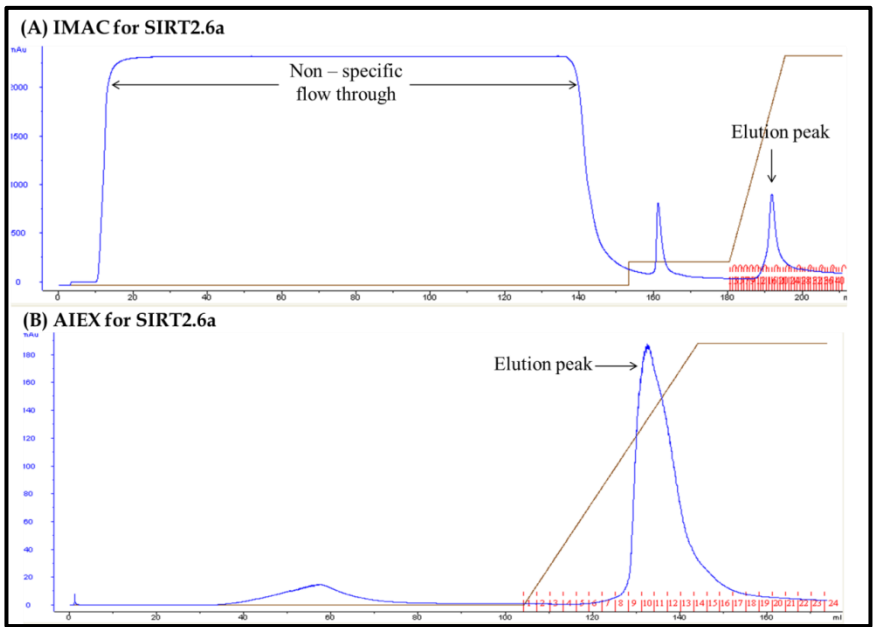


Figure 5.1. 6: SIRT2.6a purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT2.6a followed by (B) AIEX using a manually packed column using Q sepharose beads. The elution peaks are indicated in the figure.

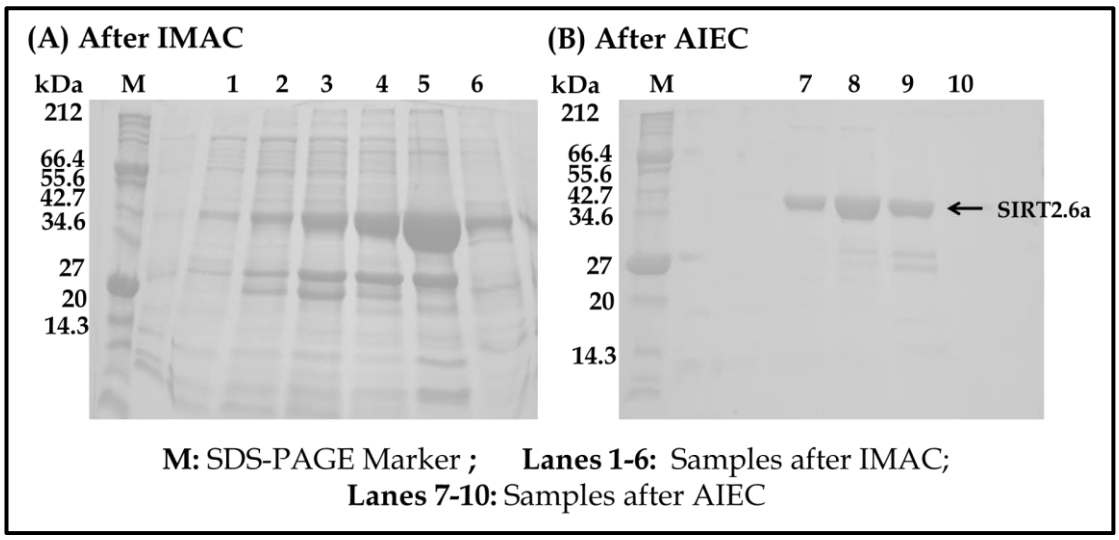


Figure 5.1. 7: Example of SIRT2.6a purification analysed by SDS-PAGE.

SIRT2.6a purification analysed by SDS-PAGE (A) samples from IMAC (lanes 1-6) followed by (B) SEC (lanes 7-10). The molecular weights of the SDS-PAGE marker (M) are labelled.

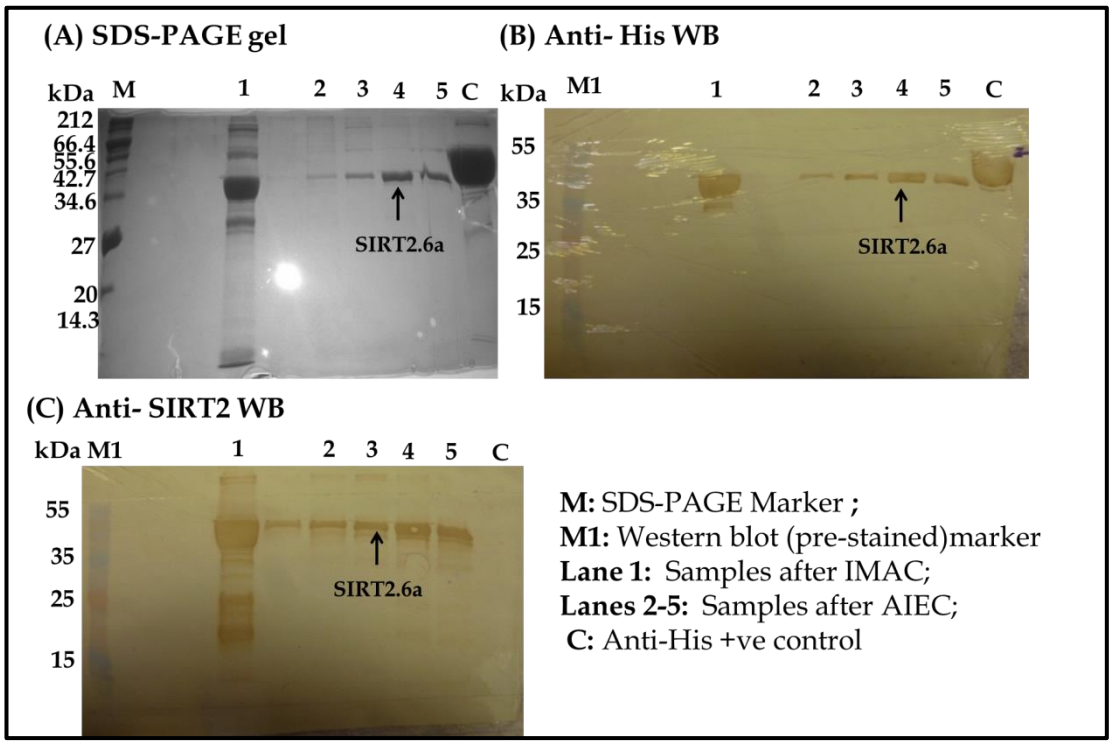


Figure 5.1. 8: Western blot analysis of SIRT2.6a.

Samples of SIRT2.6a analysed by western blotting (WB) (A) samples from IMAC (lane 1) followed by (B) SEC (lanes 2-5). The molecular weights of the SDS-PAGE marker (M) and western blot marker (M1) are labelled.

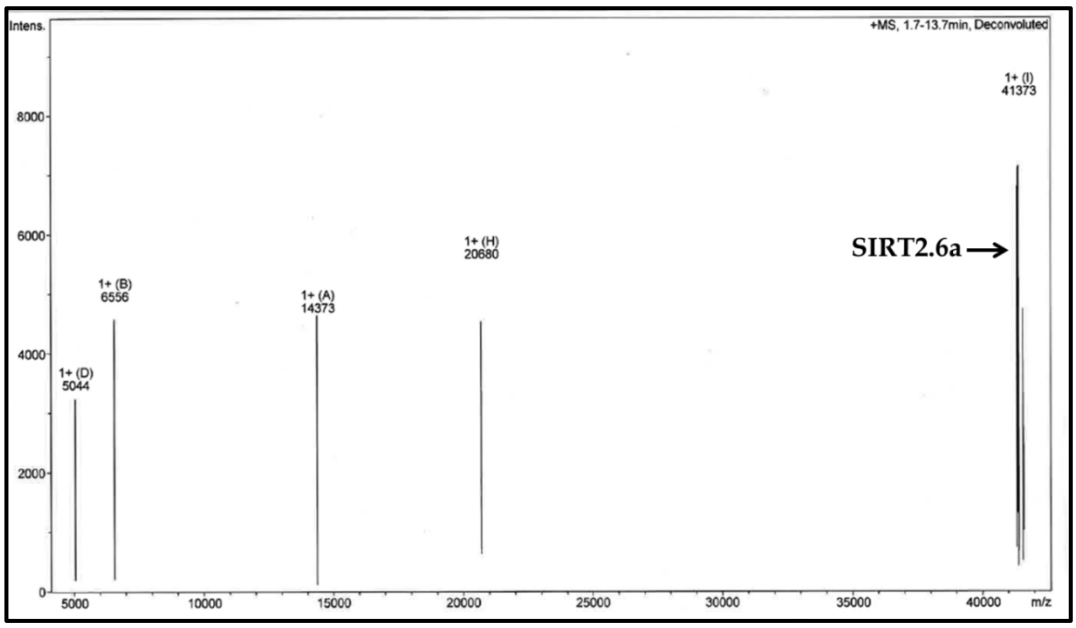


Figure 5.1. 9: Mass spectrometric analysis of SIRT2.6a.

Purified sample of SIRT2.6 analysed by mass spectrometric analysis. The peak corresponding to SIRT2.6a (41.6 kDa) is indicated.

As crystallisation of SIRT2 using SIRT2.6 was challenging (section 5.2.3.5), additional constructs SIRT2.1, SIRT2.4 and SIRT2.7 with variations in covalently linked affinity tags and truncations were expressed and purified for further crystallisation trials.

5.1.3.2. SIRT2 constructs SIRT 2.1 and SIRT2.4

The SIRT2.1 construct is a truncated at the N- and C-terminals compared to native SIRT2 and comprises residues 34-356. The recombinant gene was cloned into the pTriEx-4 Ek/LIC vector in order to facilitate the expression with a covalently linked N-terminal His and S-tag. The expressed protein is expected to have a theoretical mass of 41.5 kDa.

The SIRT2.4 construct is also truncated at the N- and C-terminals and comprises residues 34-356. The recombinant plasmid was cloned into pET-22b vector and the expressed protein is expected to have a theoretical mass of 37.6 kDa. This vector facilitates the expression with a covalently linked C-terminal His-tag.

After verification of the DNA sequences for both recombinant plasmids, SIRT2.1 was transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells while SIRT2.4 was successfully transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells. Positive results for forward and reverse DNA sequencing were obtained. The alignment of the translated DNA sequence of SIRT2.1 and SIRT2.4 with SIRT2 are given in appendix E1.3-E1.6.

Based on the expression trials for SIRT2.1 and SIRT2.4 in LB and TB media at 37 ^{o}C, LB media was selected for the large scale expression of SIRT2.1 and TB media was selected for the large scale expression of SIRT2.4 [Figure 5.1.10].

Purification of both proteins was succesfully carried out using IMAC followed by SEC [Figures 5.1.11-5.1.14]. The identity of the final purified protein was confirmed by western blot analysis using anti-His Ab and anti-SIRT2 Ab [Figure 5.1.15]. A yield of 10-15 mg of SIRT2.1 and 15-17 mg of SIRT2.4 was obtained per litre culture purified using this method. The purified protein were used for crystallisation screening (section 5.1.3.4) and SIRT2 activity assays (section 5.1.3.4).

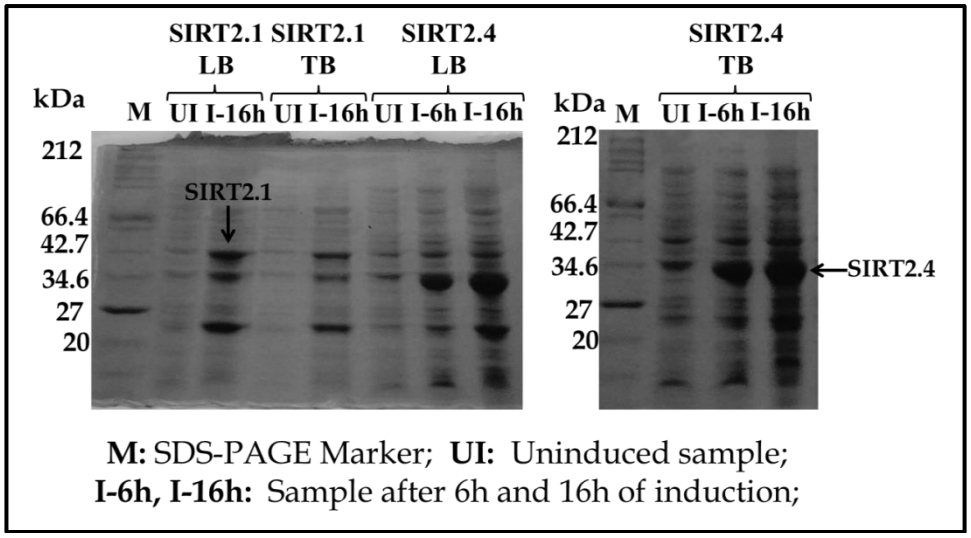


Figure 5.1. 10: Example of SIRT2.1 and SIRT2.4 expression trials analysed by SDS-PAGE.

Expression trials for SIRT2.1 and SIRT2.4 analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

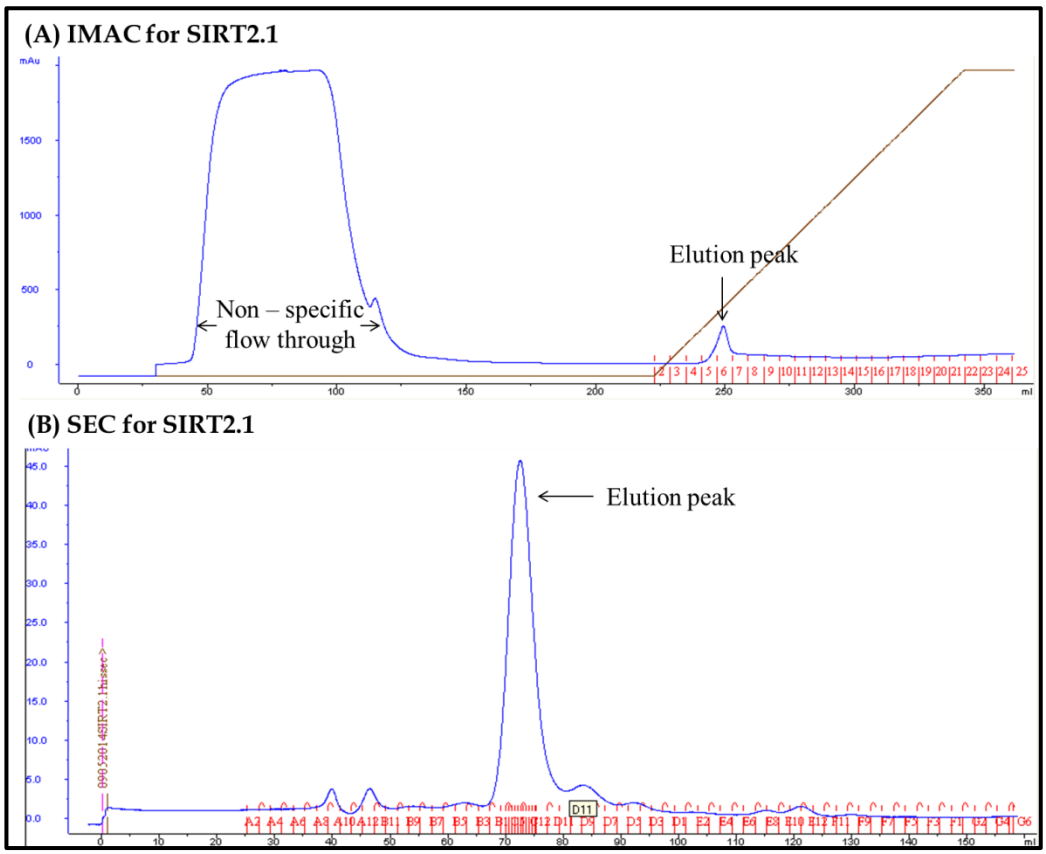


Figure 5.1. 11: SIRT2.1 purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT2.1 followed by (B) SEC using a superdex 200 column. Blue line represents the absorbance at A_{280} and the brown line represents the percentage of elution buffer.

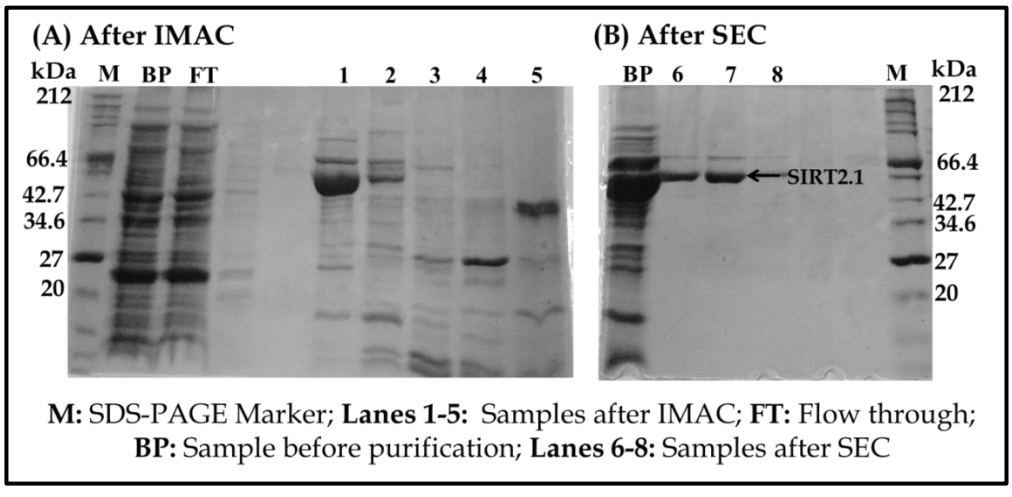


Figure 5.1. 12: Example of SDS-PAGE for SIRT2.1 purification.

SIRT2.1 purification analysed by SDS-PAGE (A) samples from IMAC (lanes 1-5) followed by (B) SEC (lanes 6-8). The molecular weights of the SDS-PAGE marker (M) are labelled.

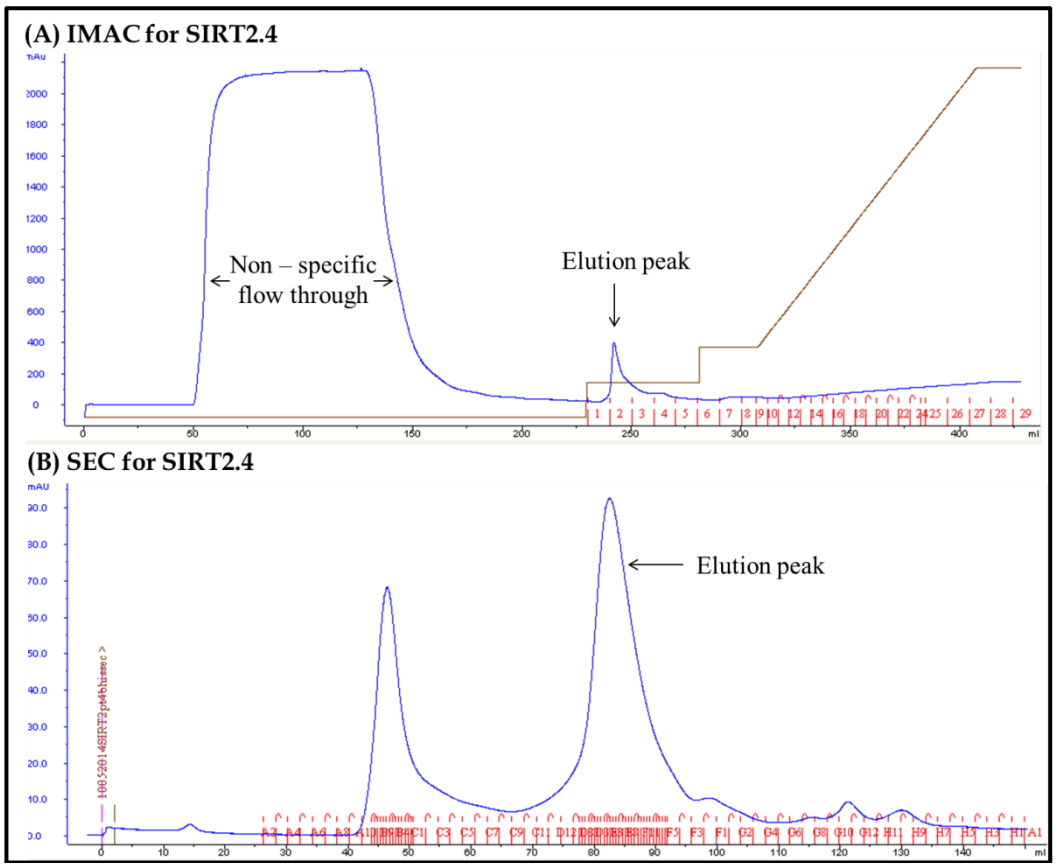


Figure 5.1. 13: SIRT2.4 purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT2.4 followed by (B) SEC using a superdex 200 column. Blue line represents the absorbance at A_{280} and the brown line represents the percentage of elution buffer.

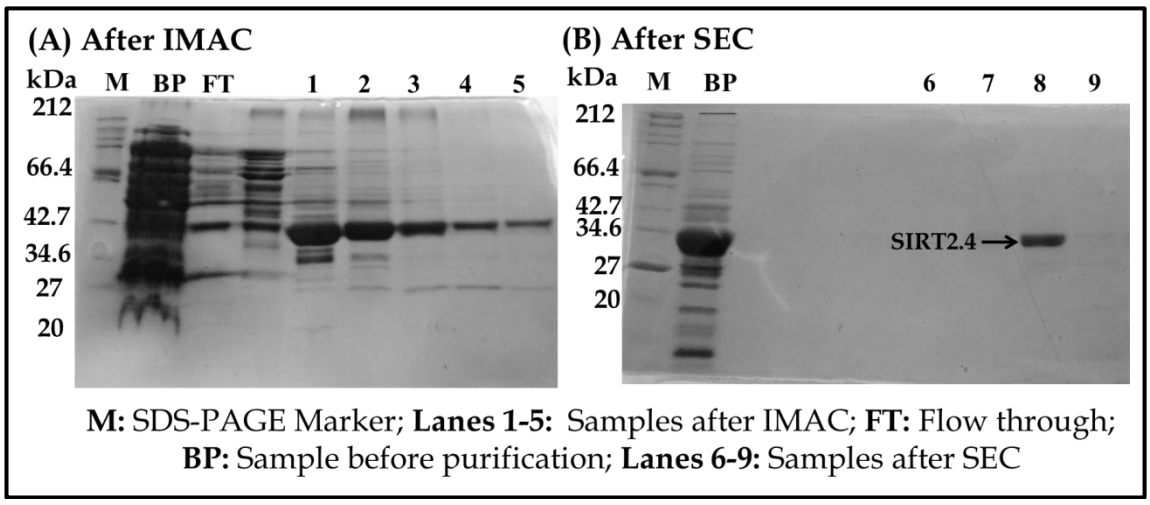


Figure 5.1. 14: Example of SIRT2.4 purification analysed by SDS-PAGE.

SIRT2.4 purification analysed by SDS-PAGE (A) samples from IMAC (lanes 1-5) followed by (B) SEC (lanes 6-9). The molecular weights of the SDS-PAGE marker (M) are labelled.

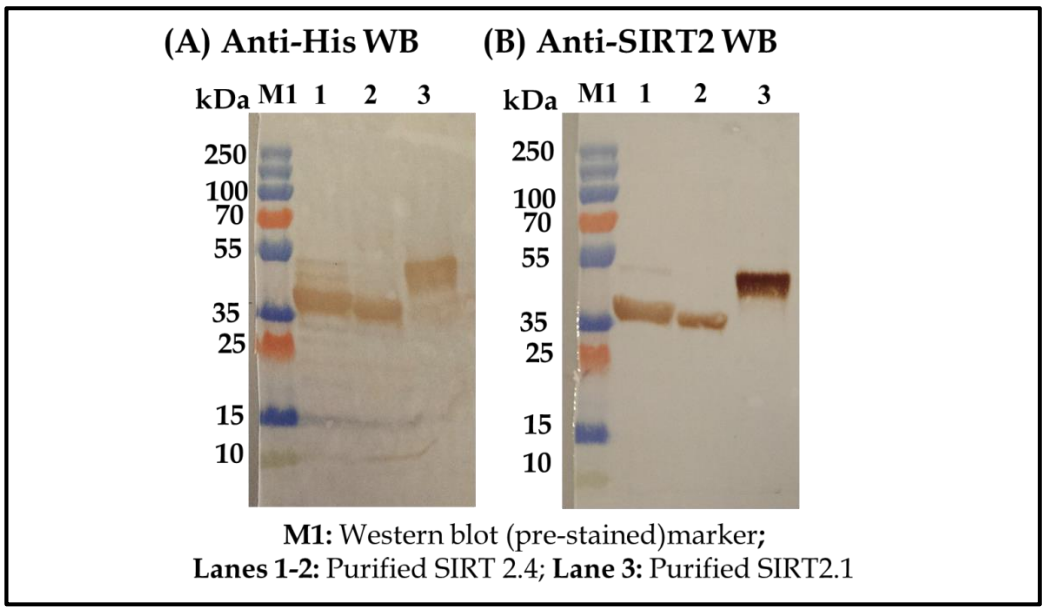


Figure 5.1. 15: Western blot analysis of SIRT2.1 and SIRT2.4.

Purified samples of SIRT2.1 and SIRT2.4 analysed by western blotting (WB) using (A)anti-His Ab (B) anti-SIRT2 Ab. The molecular weights of the western blot marker (M1) are labelled.

5.1.3.3. SIRT2 construct SIRT 2.7

The recombinant gene for construct SIRT2.7 was cloned in a pGEx-6P-1 vector to facilitate the expression of SIRT2 with a covalently linked GST-tag to aid protein expression. The primers *SIRT2.7fwd* and *SIRT2.7rev* were used for PCR amplification of the desired region of the *SIRT2.6* expression vector. A band of ~1200 bp was observed on analysis of the sample on an agarose gel [Figure 5.1.16].The resulting gene of interest was used for annealing with the PCR Blunt vector and was successfully transformed into *E. coli* TOP10 cells for DNA extraction and sequencing. Positive results for forward DNA sequencing was obtained. The alignment of the translated DNA sequence of SIRT2.7 in the recombinant bunt vector with SIRT2 is given in appendix E1.7.

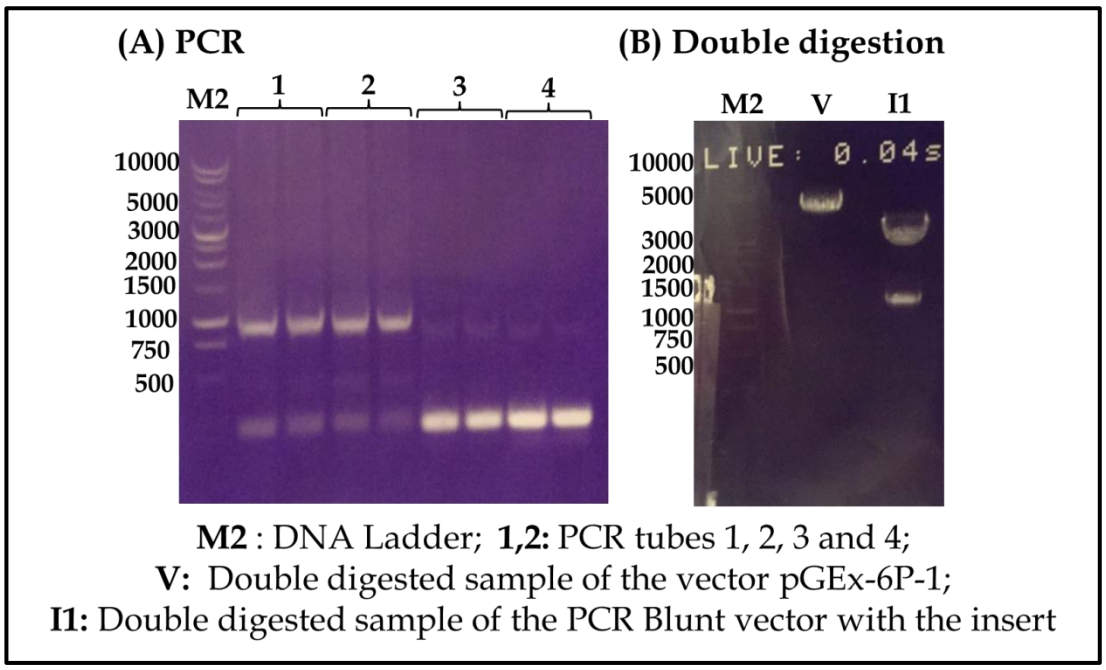


Figure 5.1. 16: SIRT2.7. cloning.

(A) Agarose gel of the PCR product for SIRT2.7. (B) Double digestion for SIRT2.7. The base pairs of the DNA ladder (M2) are labelled.

On verification of the DNA sequence, successful double digestion of the pGEx-6P-1 vector and the SIRT2.7 PCR blunt vector with *EcoRI* and *BaHI* restriction enzymes was carried out [Figure 5.1.16]. A drop out at ~1200 bp for the recombinant blunt vector and ~4000 bp for pGEx-6P-1 was observed. These bands were extracted and the drop outs were ligated to the double digested pEx-6P-1 vector. The ligated recombinant pGEx-6P-1 plasmid was transformed into TOP10 cells and the DNA was extracted for sequencing. Positive results for forward DNA sequencing was obtained. The alignment of the translated DNA sequence of SIRT2.7 with SIRT2 is given in appendix E1.8. Following verification of the DNA sequence, the plasmid was transformed into *E. coli* ArcticExpress (DE3) RIL cells to enable soluble expression of the protein at low temperatures. Expression of the protein during subsequent expression trails were analysed by SDS-PAGE and confirmed using anti-GST Ab for western blot analysis [Figure 5.1.17]. To avoid leaky expression, 10 mM glucose was added to the LB media during large scale expression trials [Figure 5.1.17].

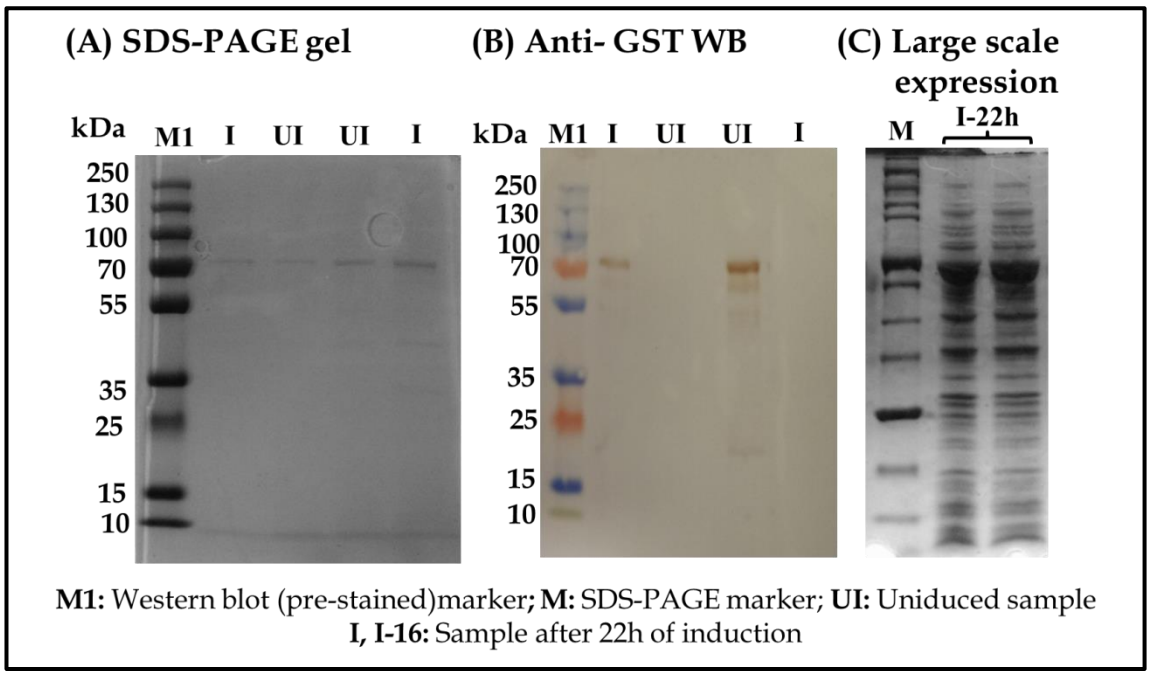


Figure 5.1. 17: Example gel for SDS-PAGE analysis for expression of SIRT2.7.

(A-B) Samples from SIRT2.7 expression trials in LB (C) samples from large scale expression of SIRT2.7. The molecular weights of the SDS-PAGE marker (M) are labelled.

Initial purification of SIRT2.7 was attained using GST-AC [Figure 5.1.18 and 5.1.19]. A western blot analysis using anti-GST Ab and anti-SIRT2 Ab confirmed the identity of the protein [Figure 5.1.20]. The purified protein was treated with PreScission protease for GST-tag cleavage and the untagged SIRT2 was separated from GST using SEC [Figure 5.1.18 and 5.1.19]. Tryptic digest of in-gel protein for mass spectrometric analysis was performed on the final purified protein to confirm the identity of the protein (~36kDa). A yield of ~5-7 mg of untagged SIRT2 per litre of cell culture was obtained from each purification. The purified protein was used in the SIRT2 activity assay and crystallisation experiments which are discussed in sections 5.1.3.4 and 5.1.3.5.

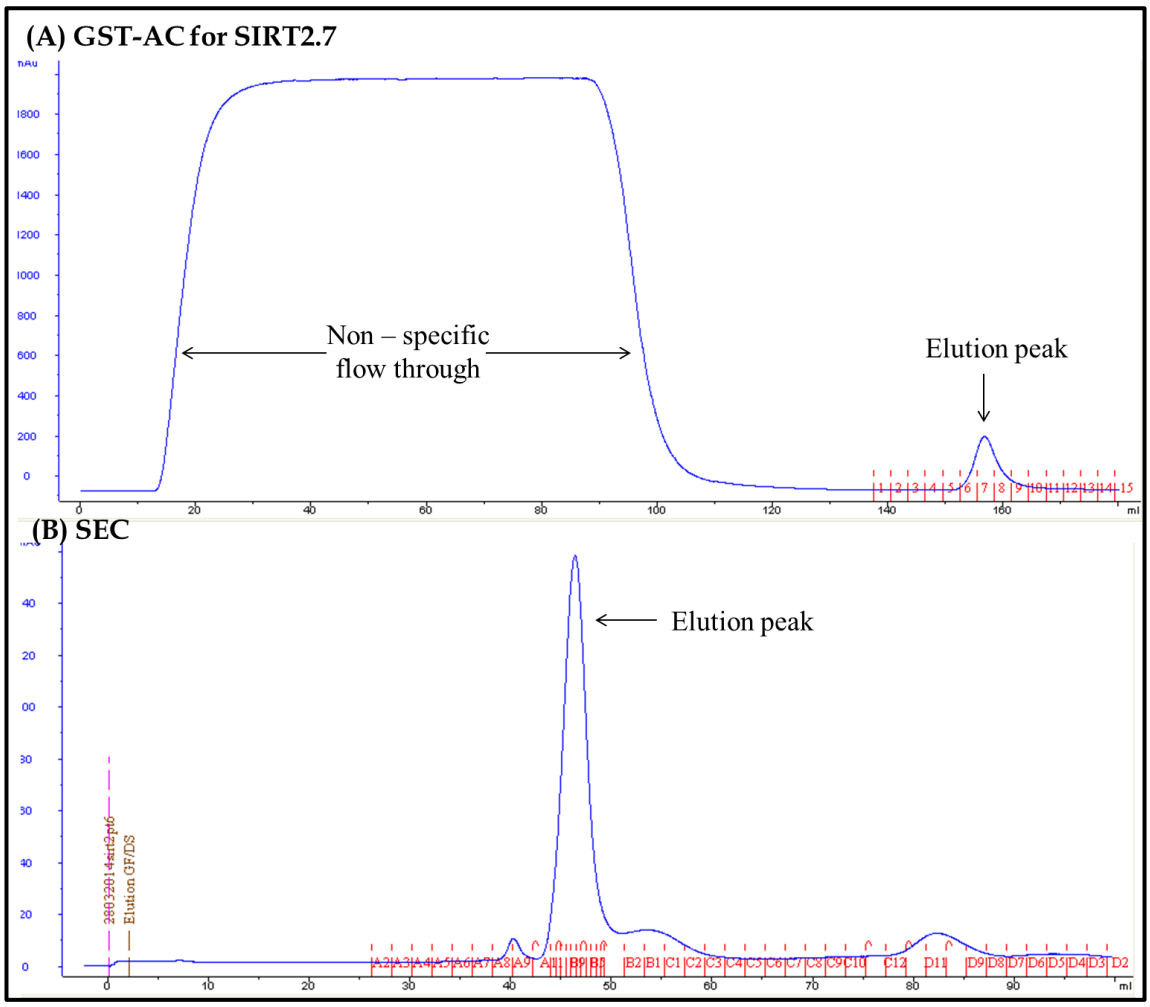


Figure 5.1. 18: SIRT2.7 purification.

Example chromatogram of (A) SIRT2.7 purification utilising a GSTrap FF column for GST-AC followed by PreScission protease treatment and (B) SEC using a superdex 200 column. Blue line represents the absorbance at A_{280}.

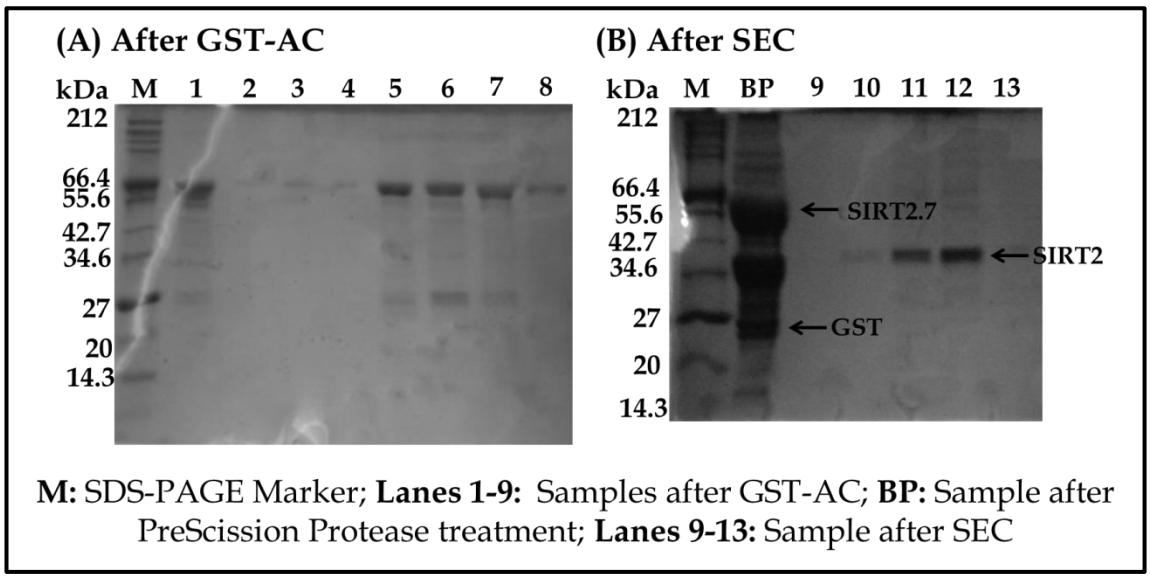


Figure 5.1. 19: Example of purification of SIRT2.7 analysed by SDS-PAGE.

SIRT2.7 purification analysed by SDS-PAGE (A) samples from GST-AC (lanes 1-8) followed by (B) PreScission protease (lane BP) and SEC (lanes 9-10). The molecular weights of the SDS-PAGE marker (M) are labelled.

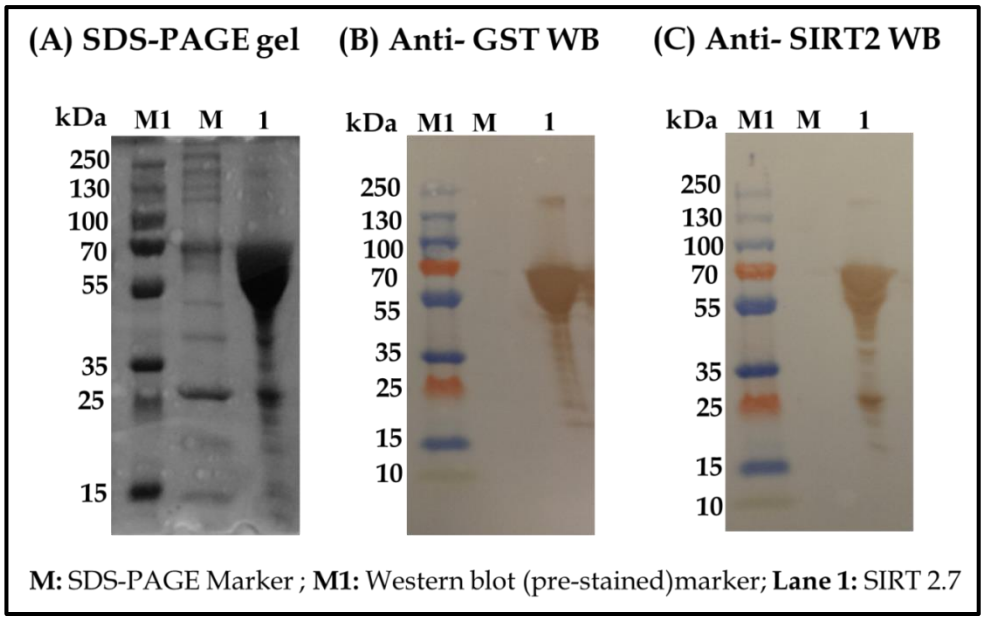
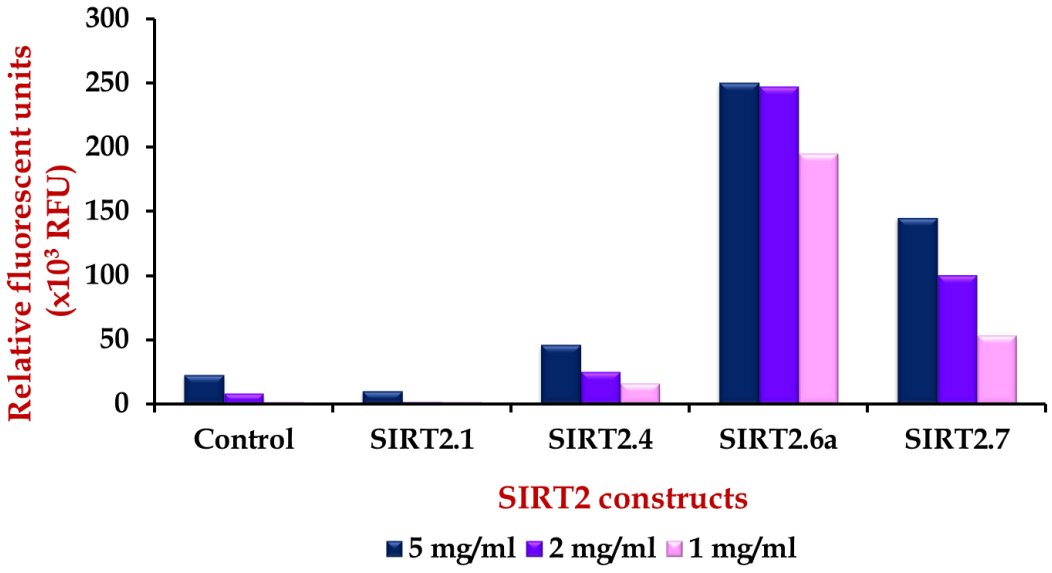


Figure 5.1. 20: Western blot analysis of SIRT2.7 after GST-AC.

SIRT2.7 analysed by western blotting (WB) using (B) anti-GST Ab (C) anti-SIRT2 Ab. The molecular weights of the SDS-PAGE marker (M) and western blot marker (M1) are labelled.

5.2.3.4. SIRT2 activity Assay

The NAD^+ binding activity of purified recombinant SIRT2 constructs SIRT2.1, SIRT2.4, SIRT2.6 and SIRT2.7 were confirmed by performing *Fluor-de-lys*®-SIRT2 activity assay. The purified protein at three different concentrations were tested of which SIRT2.6a showed the highest activity at each concentration [Figure 5.1.21]. SIRT2.1 activity was observed to be lower than the control while SIRT2.4 and SIRT2.7 along with SIRT2.6s was much higher.



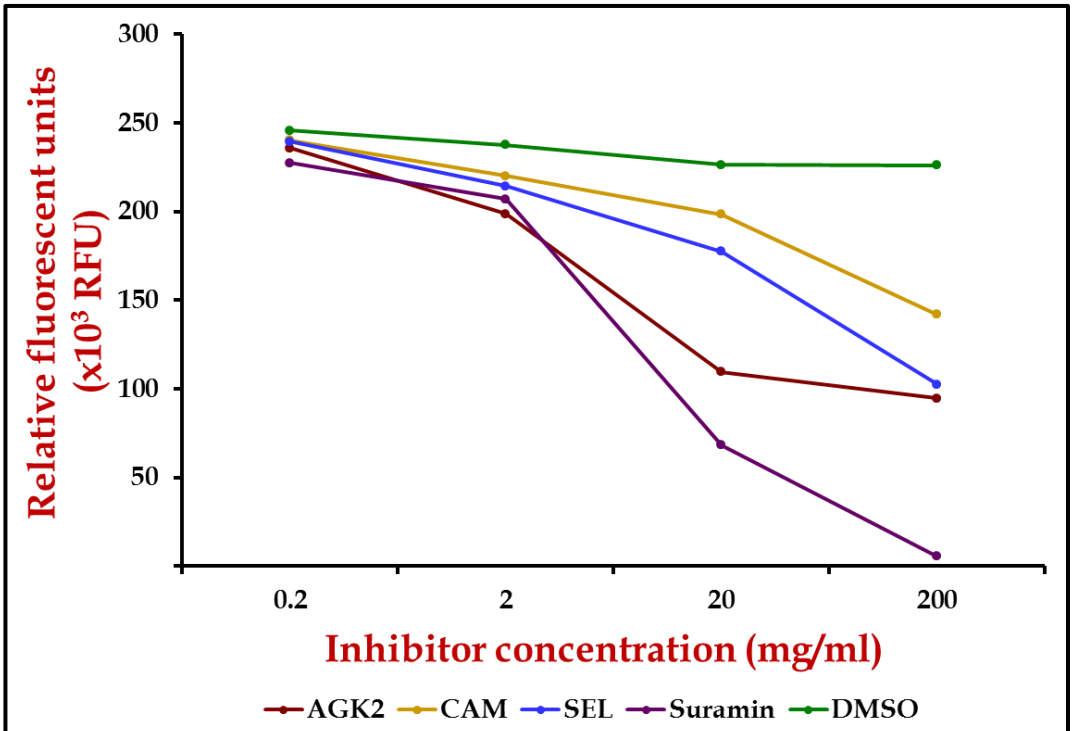
	SIRT2 activity at (RFU)		
Test sample	5 mg/ml	2 mg/ml	1 mg/ml
Control	22,997	8,411	1,040
SIRT2.1	10,441	2,041	2,041
SIRT2.4	46,312	25,411	16,383
SIRT2.6a	249,971	247,309	194,939
SIRT2.7	145,188	100,452	53,505

Figure 5.1. 21: *Fluor-de-lys*®-SIRT2 activity assay on recombinant SIRT2 constructs.

SIRT2 activity assay performed using SIRT2.1, SIRT2.4, SIRT2.6a and SIRT2.7. Recombinant SIRT2 was used as a control.

As SIRT2.6a showed the highest activity, inhibitor activity assays were performed using this protein to confirm that it recognises the inhibitor. Inhibitors, AGK2, CAM, SEL and suramin were used for this purpose. A DMSO control was also tested as AGK2, CAM and SEL were prepared in DMSO. The

control showed minimal contribution to the inhibitory activity of SIRT2.6a. The assay revealed that suramin shows the highest inhibitory activity compared to other compounds [Figure 5.1.22].



	Inhibition at (RFU)			
Inhibitor	0.2 mg/ml	2 mg/ml	20 mg/ml	200 mg/ml
AGK2	235,840	198,677	109,687	94,667
CAM	240,168	220,098	198,447	141,967
SEL	239,406	214,414	177,638	102,550
Suramin	227,431	207,143	68,412	5,570
DMSO (control)	245,832	237,571	226,381	226,235

Figure 5.1. 22: *Fluor-de-lys*®-SIRT2 inhibition activity assay on recombinant SIRT2.6a.

Inhibition of SIRT2 activity analysed using inhibitors AGK2, CAM, SEL and suramin. DMSO was used as a control as the inhibiotrs were dissolved in DMSO.

5.2.3.5. Crystallisation trials for SIRT2

Finding the winning combination that supports crystal growth often begins with experiments deigned on a coarse grid aiming to seek important experimental variables that can be explored in greater detail.

Crystallisation trials for SIRT2 started with setting up of sparse-matrix commercial conditions (Section 3.7.2). In an attempt to crystallise SIRT2, purified protein from constructs SIRT2.1, SIRT2.4, SIRT2.6 and SIRT2.7 were used for setting up automated preliminary crystallisation screening. All 12 commercial crystallisation screens (12 screens x 96 conditions each x 3 ratios x 3 temperatures x 4 protein constructs x 5 concentrations) were set up to for crystallisation of native SIRT2. The vapour diffusion sitting drop method was the preferred crystallisation technique.

A variety of conditions were tried at different temperatures (4 ^{o}C, 16 ^{o}C and 21 ^{o}C), with drops containing the protein solution and the reservoir solution in a ratios of 1:1, 1:2 and 1:3. Protein concentrations of 2 mg/mL, 6 mg/mL, 8 mg/mL, 10 mg/mL and 12 mg/mL were used for these experiments. It was observed that crystallisation of SIRT2 at concentrations above 8 mg/mL resulted in precipitation of the protein within 48 h of setting up of the crystallisation experiments. However, a mix of precipitation and clear drops were observed at other concentrations. It was also observed that incubation at 4 ^{o}C resulted in a decrease in precipitation of the protein at all concentrations. However, no crystals were observed at this temperature.

As single crystal growth was not observed, randomised microseed matrix screening for SIRT2.4, SIRT2.6 and SIRT2.7 was performed. For this, all drops from JCSG-*plus* HT-96 plates after 2 months of incubation were pooled and used as seeds to set up additional screening conditions in the hope that any crystalline precipitate present in the pool would aid crystal growth. MemGold HT-96 and JCSG-*plus* HT-96 showed potential crystallisation hits for which optimisation plates using microseeding were set up [Figure 5.1.23].

Among the various constructs used, crystallisation screening showed that untagged SIRT2.7 precipitated much faster than 6xHis tagged SIRT2.6. Even after 6 months of incubation, many clear drops of SIRT2.6 at 12 mg/mL incubated at 4 ^{o}C were observed. However, no signs of nucleation or crystallisation were seen.

As crystallisation of native protein did not yield desirable results, crystallisation screening of SIRT2.6 in complex with inhibitors was set up which yielded many salt crystals. These were confirmed to be salt crystals on exposure to X-ray radiation. However, potential signs of crystallisation were observed in two conditions for SIRT2.6 (8-14mg/mL) in complex with suramin (protein to ligand ration of 1:3) which was incubated for 45 days at 4 ^{o}C [Figure 5.1.22]. Recently, few of these conditions showed crystal like growth on incubation of SIRT 2.7 at 16 ^{o}C for 2 months. These included B9 [0.005 M nickel chloride hexahydrate, 0.1 M sodium acetate pH 5.5, 20% [w/v] PEG 4000], C4 (0.5 M potassium phosphate monobasic, 0.1M sodium cacodylate pH 6.5) and F12 [0.15 M potassium thiocyanate, 0.1 M Tris-HCl pH 7.5, 15% [w/v] PEG 6000] of the Clear Strategy II HT-96 screen. Further optimisation of these conditions have been set up.

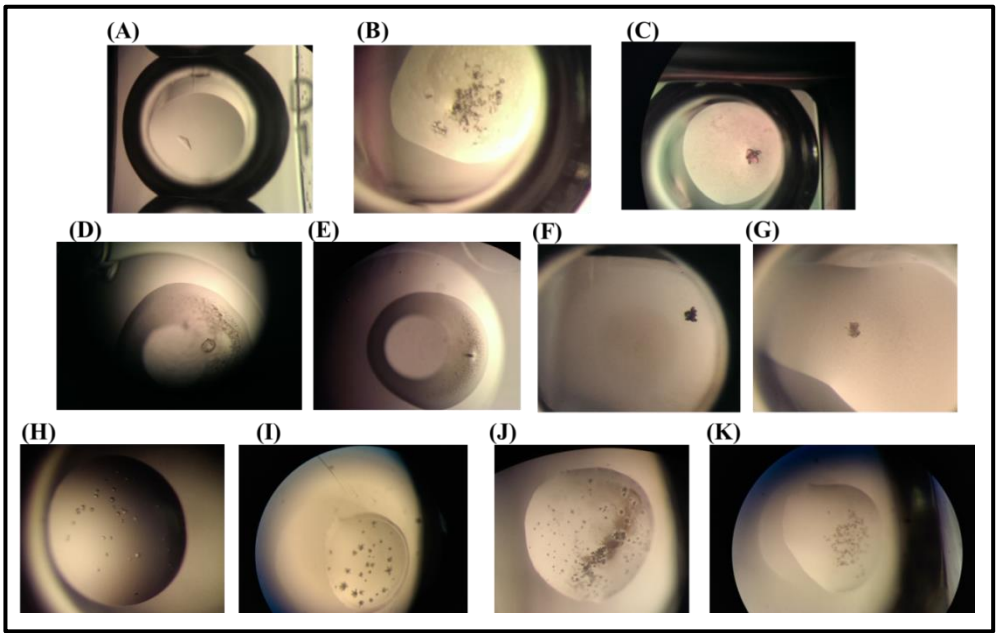


Figure 5.1. 23: Examples of crystallisation hits observed for SIRT2.

(A-C) SIRT2.6 in complex with suramin (8 mg/mL) (D-E) Native SIRT2.6 microbatch plates (F-H) SIRT2.6 + suramin (14 mg/mL) (I-K) Native SIRT2.4 (random microseed matrix screening).

As the crystals for native structure of SIRT2 (PDB: 1J8F) were grown in 24% v/v ethanol, 200 mM NaCl and 100 mM bis-tris propane HCl pH 7.0 at 4 ^{o}C, (Finnin *et al.*, 2001) crystallisation experiments based on varying these conditions were designed. Screens varying crystallisation condition for SIRT2 in complex with ADPR (20% [w/v] PEG10,000, 100 mM ammonium acetate, 100 mM bis-tris pH 5.5) (PDB: 3ZGV) (Moniot *et al.*, 2013) were also designed. However, not much success was attained from these experiments. As ethanol was one of the constituents in the condition for native SIRT2, modified microbatch plates to ensure slow diffusion of ethanol into the crystallisation drop were set up. On incubation at 16 ^{o}C for 2 months, many conditions showed crystal like growth [Figure 5.1.23]. However, exposure to X-ray radiation confirmed that these crystals were unlikely to be protein.

Although SIRT2 was expressed and purified and was shown to be biologically active, its crystallisation to date has been unsuccessful. Possible reasons for this are discussed in section 5.4.

5.2 Sirtuin 7

5.2.1. Introduction

The human Sir2 orthologue, SIRT7, is a class IV sirtuin and is the least understood member of the mammalian sirtuin family. However, it has several features suggesting its importance in human diseases such as cancer. It has poor NAD^+ dependent deacetylase activity compared to other sirtuin members. It is not yet known whether it requires activation by free fatty acids as observed for SIRT6, the closest homolog of SIRT7 and only other member of class IV sirtuin (Feldman *et al.*, 2013, Pan *et al.*, 2011) [Table 5.2.1]. Three splice variants of SIRT7 have been reported, isoforms 1-3 with 400, 183 and 320 residues, respectively (UniProt, 2014). Isoform 1 is the canonical SIRT7.

Table 5.2. 1: Overview of the results of a sequence comparison of SIRT7 with other human sirtuins.

ClustalO was used for sequence comparison (UniProt, 2014).

		Identity (%)	Identical positions
SIRT7	**SIRT1**	10.354	82
SIRT7	**SIRT2**	18.49	83
SIRT7	**SIRT3**	17.36	83
SIRT7	**SIRT4**	17.65	78
SIRT7	**SIRT5**	14.71	64
SIRT7	**SIRT6**	29.79	126

SIRT7 is a highly selective H3K18Ac deacetylase and plays a pivotal role in chromatin regulation, cellular transformation programs and tumour formation (Barber *et al.*, 2012). This is a result of depletion of the SIRT7 histone mark which is associated with highly malignant cancers and poor patient prognosis. SIRT7 is localised in the nucleolus, which is the subcellular site of ribosome assembly that increases in size and number in aggressive tumours (Hein *et al.*, 2013, Michishita *et al.*, 2005). As SIRT7 has an impact on ribosome biogenesis through multiple mechanisms, it may be playing a role in supporting the high biosynthetic and metabolic demands of cancer cells (Paredes *et al.*, 2014). These functions of SIRT7 suggest a role as a central coordinator of epigenetic and cellular homeostatic programs that support cancer progression.

The primary structure of SIRT7 is the least similar to the other member of the Sir2 family. The SIRT7 enzymatic domain shows an average similarity of 18% similarity to the highly conserved core domain of other sirtuins and 29.79% similarity to the closest sirtuin member, SIRT6 [Table 5.2.1]. The non-conserved N- and C-terminal flanking regions of SIRT7 may play a role in its relatively low deacetylase activity. These regions in other sirtuins have been shown to bind to endogenous chemical regulators (Davenport *et al.*, 2014, Sinclair and Guarente, 2014, Tennen *et al.*, 2010, Zhao *et al.*, 2003).

The highly conserved co-factor binding loop observed in other sirtuins, which is responsible for structural contacts and enzymatic activity, is replaced by a single helix containing several NAD^+ binding residues in SIRT6. The structure of SIRT7 has not yet been determined, however, predictions made based on its primary sequence comparison to other sirtuins indicate that SIRT7 also lacks this domain and is likely to adopt the same unique structural feature observed in SIRT6 (Pan *et al.*, 2011).

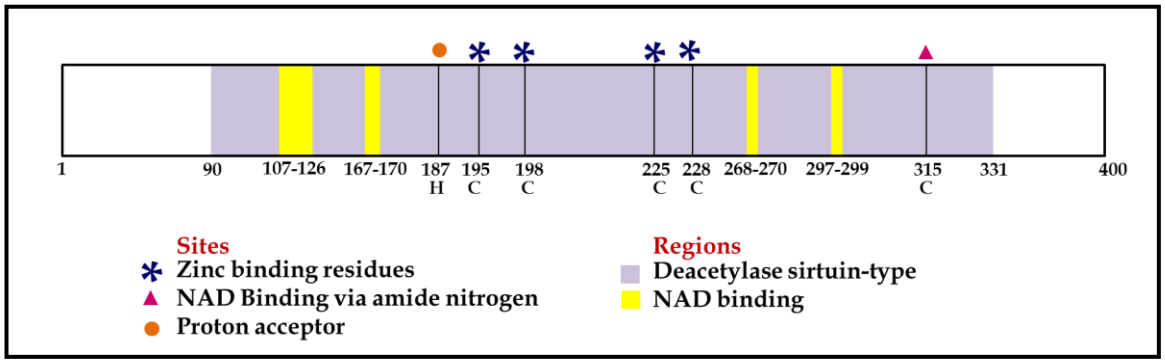


Figure 5.2. 1: Key features of the structure of SIRT7.

A schematic representation of the structure of SIRT7. The structurally important cysteine (C) residues involved in zinc binding are indicated along with the histidine (H) residue that behaves as a proton acceptor.

In contrast to other sirtuins, small-molecule modulator screening for SIRT7 has not been amenable due to its poor deacetylase activity on the substrates used for these screens. The recent identification of SIRT7 substrates has helped overcome this barrier and will aid the design of SIRT7-specific modulating compounds (Barber *et al.*, 2012, Chen *et al.*, 2013). However, deeper

understanding of SIRT7 biology at the molecular and physiological level is essential for understanding the potential therapeutic benefits and detrimental side effects of SIRT7 modulation.

To provide a clear picture of the structural contribution of SIRT7 to its elusive enzymatic activity, structural and biochemical studies of SIRT7 were initiated. In this context, the cloning, expression, purification and crystallisation of full length SIRT7 [Figure 5.2.1] was carried out. Structural understanding of SIRT7 will provide valuable insight into its physiological role.

5.2.2. Methods

5.2.2.1. Cloning, Expression and Purification

Table 5.2.2 summarises the constructs of SIRT7 that were designed for expression with covalently linked affinity tags to help expression and purification of the protein. The vector maps are provided in appendix A. The final optimised cloning, expression and protein purification protocols for each construct are given below.

Table 5.2. 2: Summary of SIRT7 constructs.

Construct	Protein insert	Vector	Tag	Mol. Wt. (kDa)
7.1	1-400	pTriEx-4 Ek/LIC	N-term His and S-tag	49.6
7.6	1-400	pET-41 Ek/LIC	N-term GST, His and S-tag	76.2
7.7	1-400	pGEx-6P-1	N-term GST tag	71.3

SIRT 7.1

The gene coding for SIRT7 was cloned into the pTriEx-4 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal poly-histidine and S-tag. This recombinant plasmid was named SIRT7.1.

The plasmid was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequencing and transformation into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells as described in Section 3.4.2. Extracted DNA was sent for sequencing to Eurofins MWG Operon, Ebersberg, Germany using T7 primer for forward sequencing. Expression trials were carried out by inoculating AI, LB and TB media with ampicillin and chloramphenicol as described in Section 3.4.1. Samples were collected at regular intervals and expression was analysed by SDS-PAGE. As insoluble expression was observed, various screens were tested to identify appropriate conditions for solubilisation of inclusion bodies and protein refolding, as described in section 3.5.2.

For large scale expression, cells were grown in LB media at 37 ^{o}C to an O.D. of 0.6–0.7 and cultures were then induced with 1 mM IPTG. Cells were harvested after 18 h of induction by centrifugation (4 ^{o}C at 12,000 g, 10 min). Induction of protein expression was analysed by SDS-PAGE.

Refolded SIRT7.1 was purified using IMAC followed by SEC. Cell pellets were resuspended in lysis buffer (50 mM Tris HCl pH 7.4) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 ^{o}C at 30,000 g, 30 min). The pellet was resuspended in 50 mM HEPES pH 7.4, 0.1% v/v Triton x100 and clarified by centrifugation (4 °C at 10,000 g, 30 min). Following the wash step, inclusion bodies were solubilised in the solubilisation buffer (50 mM Tris HCl, 500 mM NaCl, 7 M guanidine HCl, 0.15 M reduced glutathione) and the sample was left stirring overnight at 4 ^{o}C. A drop-wise rapid dilution method was employed for refolding in the refolding buffer (50 mM Tris HCl pH 7.4, 0.1 % [w/v] PEG6000, 200 mM NaCl), which was subsequently left undisturbed overnight. The preparation was clarified by centrifugation (4 °C at 12,000 g, 15 min). The supernatant was loaded on a 5 mL HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl, 15 mM imidazole). The column was washed extensively and a gradient elution of 0-100% buffer B (50 mM Tris HCl pH 7.4, 200 mM NaCl, 500 mM imidazole) over 50 mL was carried out. The eluted fraction was dialysed overnight in wash buffer at 4 °C and loaded on a gel filtration column (16/600 Superdex 200 pg; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl). The eluted samples were analysed by SDS-PAGE. Western blot analysis using anti-His Ab (Sigma-Aldrich Co., Kent, U.K.) and mass spectrometric analysis of the samples were carried out to confirm the identity of the protein.

SIRT7.1a

Alternatively, SIRT7.1 was purified by another method (referred to as SIRT7.1a). SIRT7.1a (pI 9.65) was purified by cation exchange chromatography (CIEX). The refolded protein was loaded on a manually packed 10 mL sulphopropyl (SP) sepharose (Sigma Aldrich, MO, USA) column pre-equilibrated with wash buffer (50 mM Tris HCl pH 7.4). The column was washed extensively and a gradient elution of 0-100% buffer B (50 mM Tris HCl pH 7.4, 1 M NaCl) over 100 mL was carried out. The eluted samples were analysed by SDS-PAGE.

SIRT7.6

The gene coding SIRT7 was cloned into pET-41 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal GST-tag. This recombinant plasmid was named SIRT7.6.

The recombinant plasmid *SIRT7.1* was used for PCR amplification as described in Section 3.4.2. The primers used for PCR amplification are given in table 5.2.3, while the components of the reaction mixtures and the parameters for thermal cycling are given in tables 5.2.4 and 5.2.5, respectively.

Table 5.2. 3: Primers used for PCR reaction of SIRT7.6

Primer	**Sequence (5′-3′)**
SIRT7.6fwd	GAC GAC GAC AAG ATG GCA GCC GGG CTG AGC CGC
SIRT7.6rev	GAG GAG AAG CCC GGT TAC GTC ACT TTC TTC CTC CCT GTG CG

Table 5.2. 4: PCR reaction for SIRT7.6

Reaction	Tube 1 (µL)	Tube 2 (µL)	Tube 3 (µL)
DNA template (SIRT7.1)(100ng/µL)	2	2	2
Primer (*SIRT7.6fwd*) (10 µM)	1.5	1.5	1.5
Primer (*SIRT7.6rev)* (10 µM)	1.5	1.5	1.5
KOD polymerase (1U/µL)	1	1	1
10x Buffer	5	5	5
dNTPs (2mM each)	5	5	5
$MgSO_4$ (25 mM)	3	3	3
Betaine	-	10	-
DMSO	-	-	2
Water (q.s. 50 µL)	31	21	29

Table 5.2. 5: PCR thermal cycle for SIRT7.6

Cycle	Temperature (°C)	Time (min)
Initial denaturation	94	2
Denaturation	94	0.5
Annealing	62	1
Extension	72	1
Back to step 2 (Denaturation) for 35x		
Final extension	72	10
Hold at 4°C		

The PCR products were analysed by running the samples on an agarose gel as described in section 3.4.1. The final purified product was ligated into the pET-41Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal GST, His and S-tag. The recombinant vector was transformed into *E. coli* TOP10 cells and the DNA was extracted as described in section 3.4.2. The sequence verified plasmid was then transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells and glycerol stocks were prepared as described in Section 3.4.2. Expression trials were carried out by inoculating in LB, TB, modified TB, HIS, ZXY7, SOC, glucose M9Y, superbroth, TYGPN and NZCYM media supplemented with kanamycin and chloramphenicol as described in Section 3.4.1. Composition of

the various media is given in appendix B. Samples were collected at regular intervals and protein expression was analysed by SDS-PAGE.

SIRT7.6a

The recombinant plasmid was transformed into *E. coli* ArcticExpress (DE3) RIL cells and glycerol stocks were prepared as described in Section 3.4.2. This transformant was named SIRT7.6a.

Expression trials in LB media were carried out as per manufacturer's instructions for *E. coli* ArcticExpress (DE3) RIL cells. For large scale expressions, overnight cultures were grown in LB media supplemented with gentamycin and kanamycin. Large cultures were grown by inoculating modified LB media (+10 mM $MgCl_2$ and 10 mM glucose) containing no antibiotics with overnight cultures. These large scale cultures were grown for 6 h at 30 ^{o}C before being induced with 1 mM IPTG. The cells were then grown at 10 ^{o}C overnight. Cells were harvested after 18 h of induction by centrifugation (4 °C at 12,000 g, 10 min). Induction of protein expression was analysed by SDS-PAGE.

SIRT7.6a was purified using GST-AC followed by GST-tag cleavage and SEC. Cell pellets were resuspended in PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and were lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 30,000 g, 30 min). The supernatant was loaded on a 5 mL GSTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with PSB. The column was washed extensively with PBS and a single step elution (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione) was carried out. The eluted fraction was concentrated and desalted using a HiPrep 26/10 desalting column (GE Healthcare, Uppsala, Sweden). The eluate was concentrated and loaded on a gel filtration column (16/600 Superdex 200 pg; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl). The eluted samples were analysed by SDS-PAGE. Western blot analysis using anti-GST Ab (Sigma-Aldrich Co., Kent, U.K.) and in–gel tryptic digest mass spectrometric analysis of the samples were carried out to confirm the identity of

the protein. GST-tag cleavage trial experiments with enterokinase (Trevigen, MD, USA) were carried out as per manufacturer's protocols.

SIRT 7.7

The gene coding SIRT7 was cloned into the pGEx-6P-1 vector (GE Healthcare, U.K.) with a cleavable N-terminal GST-tag. The recombinant plasmid *SIRT7.1* was used for PCR amplification as described in section 3.4.2. Primers were designed such that a *EcoR1* restriction site was incorporated into the DNA using the forward primer and an *Xho1* restriction site was incorporated using the reverse primer. The primers used for PCR amplification are given in table 5.2.6, while the components of the reaction mixtures and the parameters for thermal cycling are given in tables 5.2.7 and 5.2.8, respectively.

Table 5.2. 6: Primers used for PCR reaction of SIRT7.7

Primer	**Sequence (5′-3′)**
SIRT7.7fwd	GAA TTC ATG GCA GCC GGG GGT CTG
SIRT7.7rev	CTC GAG TTA CGT CAC TTT CTT CCT

Table 5.2. 7: PCR reaction for SIRT7.7

Reaction	**Tube 1 (µL)**	**Tube 2 (µL)**
DNA template (SIRT7.1) (100ng/µL)	2	2
Primer (*SIRT7.7fwd*) (10 µM)	1.5	1.5
Primer (*SIRT7.7rev)* (10 µM)	1.5	1.5
KOD polymerase (1U /µL)	1	1
10x Buffer	5	5
dNTPs (2 mM each)	5	5
$MgSO_4$ (25 mM)	2	2
Betaine	10	10
DMSO	2	-
Water (q.s. 50 µL)	20	22

Table 5.2. 8: PCR thermal cycle for SIRT7.7

Cycle	Temperature (°C)	Time (min)
Initial denaturation	95	2
Denaturation	95	0.5
Annealing	50	1
Extension	72	1
Back to step 2 (Denaturation) for 35x		
Final extension	72	10
Hold at 4°C		

The PCR products were analysed by running the samples on an agarose gel as described in section 3.4.1. The final purified product was ligated with a pCR™-Blunt vector (Zero Blunt® PCR Cloning Kit, Invitrogen, CA, USA). The recombinant vector was transformed into *E. coli* TOP10 cells and the DNA was extracted as described in section 3.4.2. Double digestion experiments to test for the presence of an insert were performed by incubating the reaction mixture [Table 5.2.9] for 1 h at 37 °C. FastDigest XhoI and EcoRI (Thermo scientific, MA, USA) were used for double digestion of the recombinant blunt vector (RBV) and pGEx-6P-1 vector (GE Healthcare, Uppsala, Sweden).

Table 5.2. 9: Double digestion set up for SIRT7.7.

	Test Double Digestion (µL)	Final Double Digestion (µL)	Vector Double Digestion (µL)
Xho1	1	4	4
EcoR1	1	4	4
RBV / (pGEx-6P-1 vector)	6	25	25
10x FastDigest Green buffer	1	5	5
Water	1 (q.s. 10 µL)	12 (q.s. 50 µL)	12 (q.s. 50 µL)

Following successful test double digests, the recombinant blunt vector was sent for sequencing to Eurofins MWG Operon, Ebersberg, Germany using the M13(-21) primer. Final double digestion of the RBV and pGEx-6P-1 vector was carried out on verification of the sequencing results [Table 5.2.9]. The Rapid DNA Ligation kit (Thermo Scientific, MA, USA) was used for ligations. Ligation of

the drop out (~1400 bp) from the RBV and the treated pGEx-6P-1 was carried out by incubating the ligation reaction (1 µL T4 DNA ligase, 1 µL treated pGEx-6P-1, 6 µL RBV drop out, 2 µL 5x rapid ligation buffer) for 1 h at 21 ^{o}C. The recombinant plasmid was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequence analysis as described in Section 3.4.2. On verification of sequence, the final recombinant plasmid was transformed into *E. coli* ArcticExpress (DE3) RIL cells and glycerol stocks were prepared as described in Section 3.4.2. Expression trials in LB media were carried out as per the manufacturer's instructions for *E. coli* ArcticExpress (DE3) RIL cells.

For large scale expressions, overnight cultures were grown in LB media with gentamycin and ampicillin. Large cultures were grown by inoculating modified LB media (+10 mM glucose) containing no antibiotics with overnight cultures. These cultures were grown for 6 h at 30 ^{o}C before being induced with 1 mM IPTG and grown overnight at 10 ^{o}C. Cells were harvested after 18 h of induction by centrifugation (4 °C at 12,000 g, 10 min). Induction of protein expression was analysed by SDS-PAGE.

SIRT7.7 was purified using GST-AC followed by GST tag cleavage and a second GST-AC to separate the cleaved protein from the tag. Cell pellets were resuspended in PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at 30,000 g, 30 min). The supernatant was loaded on a 5 mL GSTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with PBS. The column was washed extensively and a single step elution (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione) was carried out. The eluted fraction was concentrated and desalted using a HiPrep 26/10 desalting column (GE Healthcare, Uppsala, Sweden). Western blot analysis using anti-GST Ab and in–gel tryptic digest mass spectrometric analysis of the samples were carried out to confirm the identity of the protein. PreScission protease (GE Healthcare, Uppsala, Sweden) was added to the eluted protein (one unit protease cleaves 90% 100 µg of GST fusion protein) and the sample was dialysed overnight in 50 mM Tris HCl pH 8.0, 10 mM reduced

glutathione at 4 °C. The GST cleaved protein was loaded on a 1 mL GSTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM Tris HCl pH 8.0, 200 mM NaCl). SIRT7 was collected in the flow through and crystallisation experiments were set up.

5.2.2.2. SIRT7 activity assay

To detect the ADP-ribosylation activity of SIRT7, western blot analysis using Biotinylated NAD^+ (Trevigen, MD, USA) and horseradish peroxidase (HRP)-conjugated streptavidin (Thermo scientific, MA, USA) was performed. Biotinylated NAD^+ was supplied at 25 µM in water. Refolding and activity of SIRT7.1 was confirmed by this assay. SIRT7.1 samples were prepared (7µL SIRT7 at 3 mg/mL + 10 µL biotinylated NAD^+ 3 µL water) and analysed as per western blot analysis described in 3.5.3.2. However, 1% [w/v] bovine serum albumin (BSA) was used in place of milk powder throughout the experiment.

5.2.2.3. Crystallisation experiments

Crystallisation trials for SIRT7 started with sparse-matrix commercial conditions (Section 3.7.2). The vapour diffusion sitting drop method was the preferred crystallisation technique. A variety of conditions were tried at different incubation temperatures (4 °C, 16 °C and 21 °C), with drops containing the protein solution and the reservoir solution in ratios of 1:1, 1:2 and 1:3.

The protein obtained from constructs SIRT7.1, 7.1a, 7.6a and 7.7 were used for setting up expression trials as described in section 3.7.2. Initial crystallisation trials of native SIRT 7 were set up in their respective final purification buffers. Protein concentrations of 2 mg/mL, 6 mg/mL and 8 mg/mL were used for these experiments. Additional trials after dialysis of native SIRT7 into five other buffers [Table 5.2.10], were also set up.

Table 5.2. 10: Buffers used for the dialysis of SIRT7 prior to setting up crystallisation screens.

Buffer	Components
Buffer 1	50 mM Tris HCl pH 7, 200 mM NaCl
Buffer 2	100 mM Tris HCl pH 8, 150 mM NaCl
Buffer 3	50 mM Tris HCl pH 7.4, 200 mM NaCl, 0.01% v/v TCEP
Buffer 4	50 mM HEPES pH 7.4, 150 mM NaCl
Buffer 5	50 mM HEPES pH 6, 150 mM NaCl

Commercial crystallisation screens for SIRT7.6 in complex with NAD^+ and SIRT7.6 (6 mg/mL) in complex with ADPR were set up at molar ratios of 1:1, 1:2, 1:4 and 1:10 of protein to NAD^+ or ADPR. Systematic screening of hits was performed and is discussed further in results section 5.2.3.5.

5.2.3. Results

5.2.3.1. SIRT7 construct SIRT7.1

The recombinant gene for full length SIRT7 was cloned into the pTriEx-4 Ek/LIC vector in order to facilitate its expression with a covalently linked His-tag. The expressed protein is expected to have a theoretical mass of 50 kDa. Following verification of a correct DNA sequence and subsequent transformation, expression trials in LB, TB and AI media at 37 ^{o}C for construct SIRT7.1 were carried out. SIRT7 expression was observed in all three media [Figure 5.2.2]. Positive results for forward DNA sequencing was obtained. The alignment of the translated DNA sequence of SIRT7.1 with SIRT7 is given in appendix E2.1. As expression in LB media showed the lowest level of impurities, it was selected as the growth medium for further expressions. However, SIRT7 was expressed in the form of inclusion bodies as opposed to a soluble form [Figure 5.2.2].

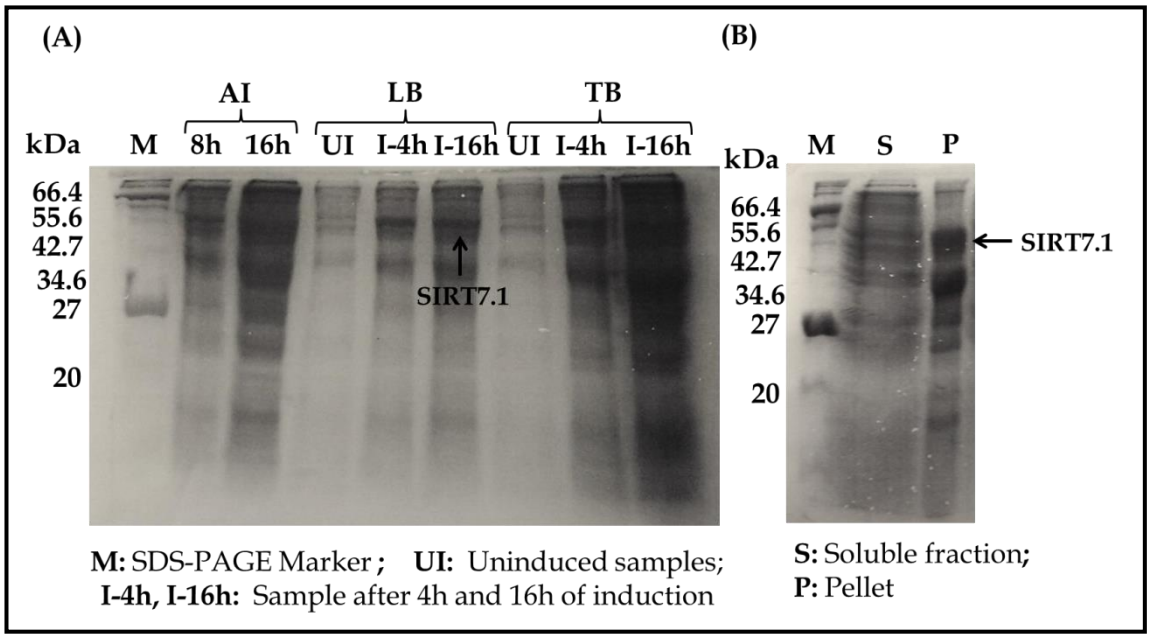


Figure 5.2. 2: SIRT7.1 expression trials analysed by SDS-PAGE.

(A) SIRT7.1 expression trial in AI, LB and TB media at 37 ^{o}C. (B) The soluble fraction and the pellet after cell disruption of the I-16h expression sample in LB media analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

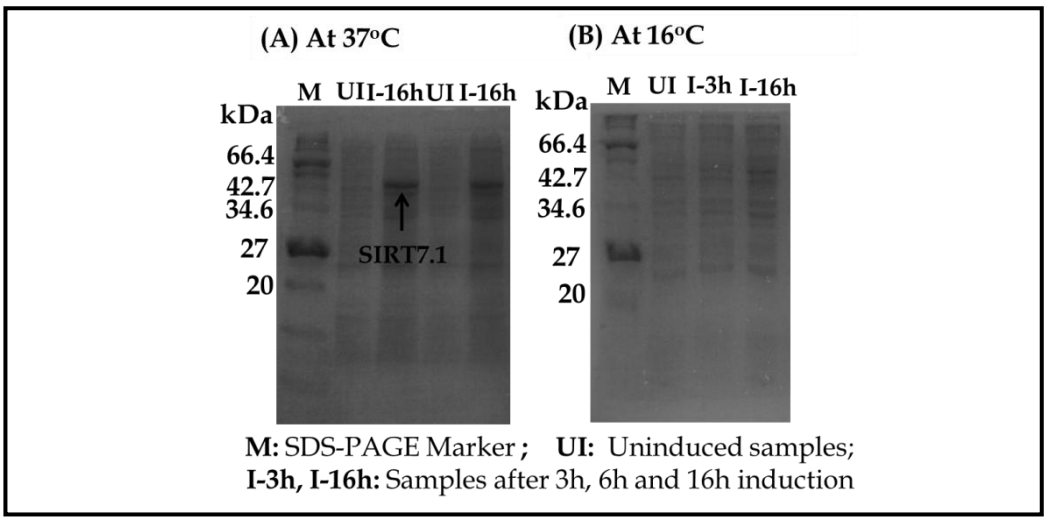


Figure 5.2. 3: SIRT7.1 expression in LB media analysed by SDS-PAGE.

SIRT7.1 expression trials analysed by SDS-PAGE. (A) Expression trial at 37 °C. (B) Expression trial at 16 °C. The molecular weights of the SDS-PAGE marker (M) are labelled.

To confirm the insoluble expression observed was not due to expression at a high temperature, expression trials at 16 °C were carried out in LB. This trial confirmed that insoluble expression at 37 °C was more favourable as no protein expression was not observed at 16 °C [Figure 5.2.3]. Solubilisation and refolding trials were performed to select the appropriate conditions for treatment of the inclusion bodies. Condition C (100 mM Tris HCl pH 8, 500 mM NaCl, 7 M guanidine HCl, 150 mM reduced glutathione) in the solubilisation trial was identified as most suitable. This was based on the protein being observed in the soluble fraction while the impurities remained in the pellet [Figure 5.2.4]. Refolding of the protein was achieved using conditions E8 (50 mM Tris HCl pH 8, 500 mM L-Arginine) and F3 (50 mM Tris HCl pH 8, 100 mM NaCl, 0.1% [w/v] PEG 6000). Successful purification of SIRT7.1 was achieved using the E8 refolding condition followed by IMAC and SEC [Figure 5.2.5 and 5.2.6]. A yield of 5-7 mg per litre of culture was observed. The identity of the sample was confirmed using mass spectrometric analysis [Figure 5.2.7] and western blot analysis using anti-His Ab [Figure 5.2.8]. Additionally, CIEX of the refolded protein (referred to as SIRT7.1a) was attempted to attain higher purity and yield of SIRT7.1 [Figure 5.2.9 and 5.2.10]. A yield of 10-15 mg per litre of culture was

obtained using this method. Crystallisation screening, which was set up using both SIRT7.1 and SIRT7.1a protein, is discussed in section 5.2.3.4.

A major concern for proteins expressed as inclusion bodies is correct refolding of the protein after denaturation. Correct refolding of the protein is crucial for the protein to be active. To confirm the retention of activity and correct refolding of SIRT7.1a, SIRT7 activity assay was performed which is discussed in section 5.2.3.4. This purified protein was also used to set up crystallisation experiments discussed in section 5.2.3.5.

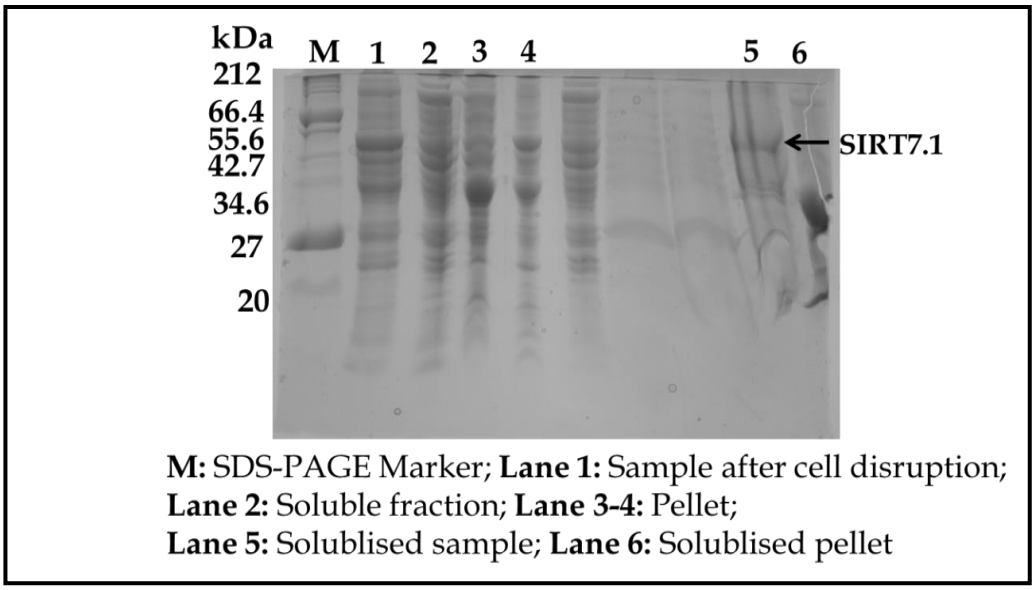


Figure 5.2. 4: Example of solubilisation and refolding of SIRT7.1 analysed by SDS-PAGE.

The molecular weights of the SDS-PAGE marker (M) are labelled.

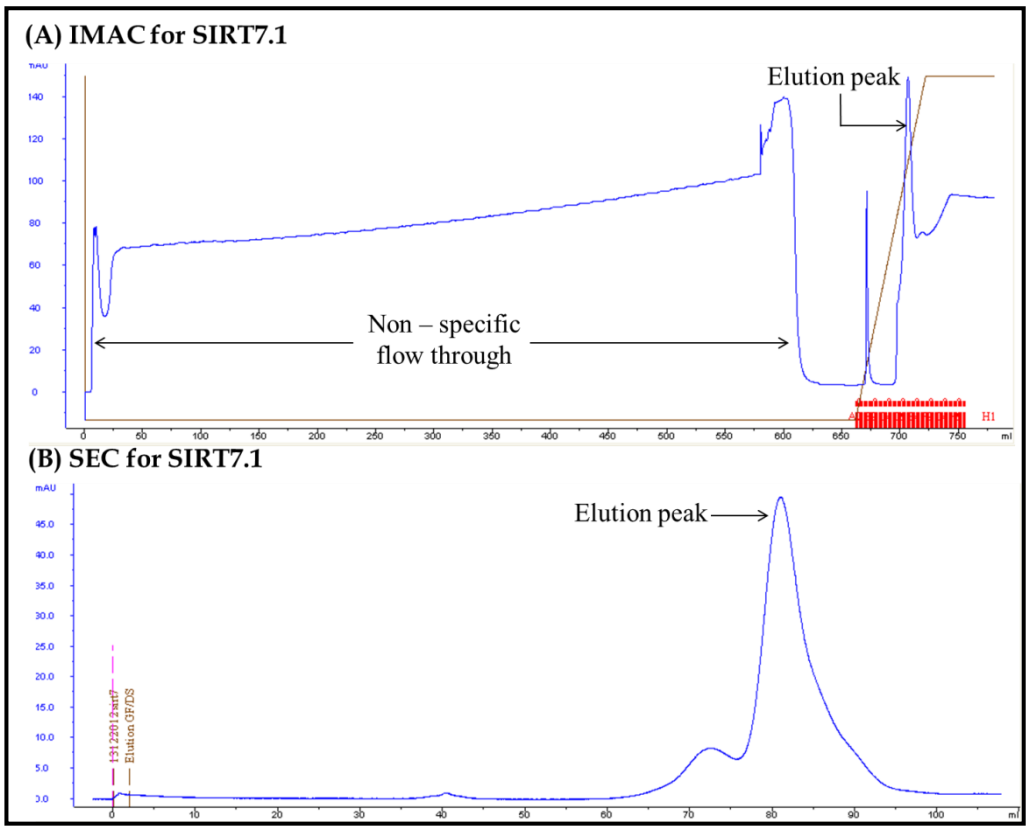


Figure 5.2. 5: SIRT7.1 purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT7.1 followed by (B) SEC using a superdex 200 column. Blue line represents the absorbance at A_{280} and the brown line represents the percentage of elution buffer.

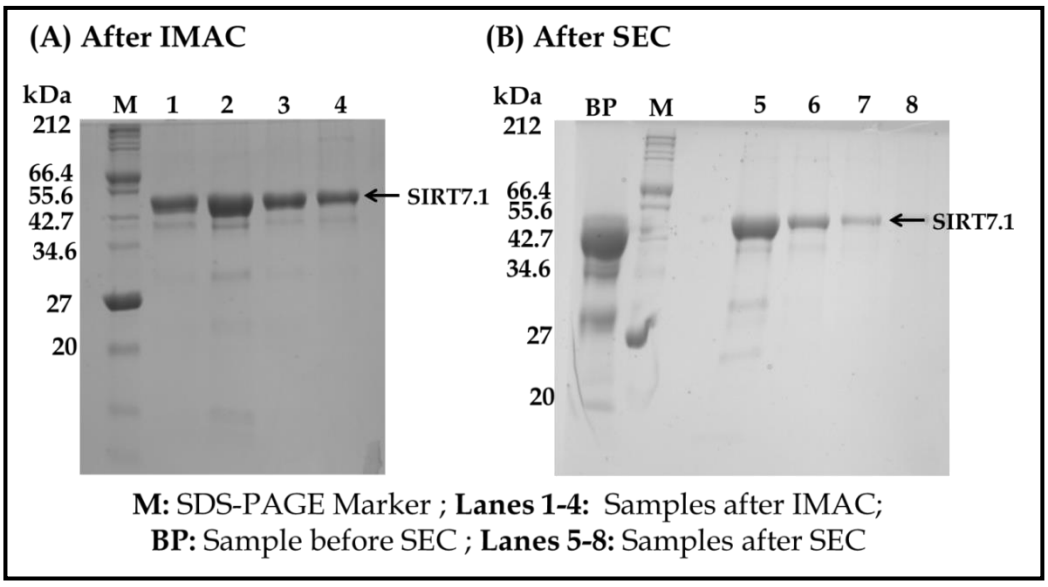


Figure 5.2. 6: Example of SIRT7.1 purification analysed by SDS-PAGE.

SIRT7.1 purification analysed by SDS-PAGE (A) samples from IMAC (lanes 1-4) followed by (B) SEC (lanes 5-8). The molecular weights of the SDS-PAGE marker (M) are labelled.

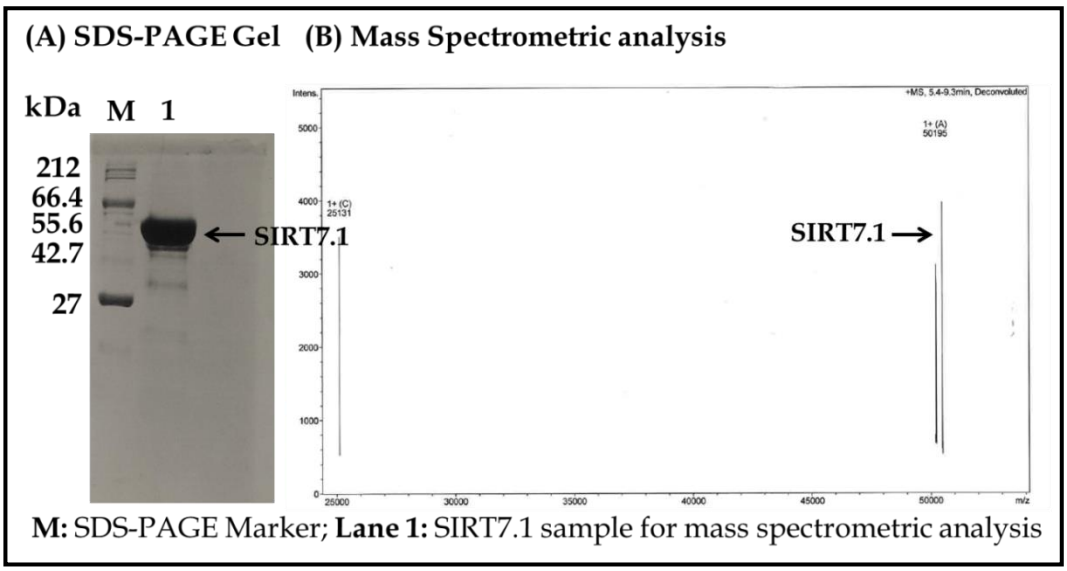


Figure 5.2. 7: Mass spectrometric analysis of SIRT7.1.

Purified sample of SIRT7.1 analysed by mass spectrometric analysis. The peak corresponding to SIRT7.1 (49.6 kDa) is indicated.

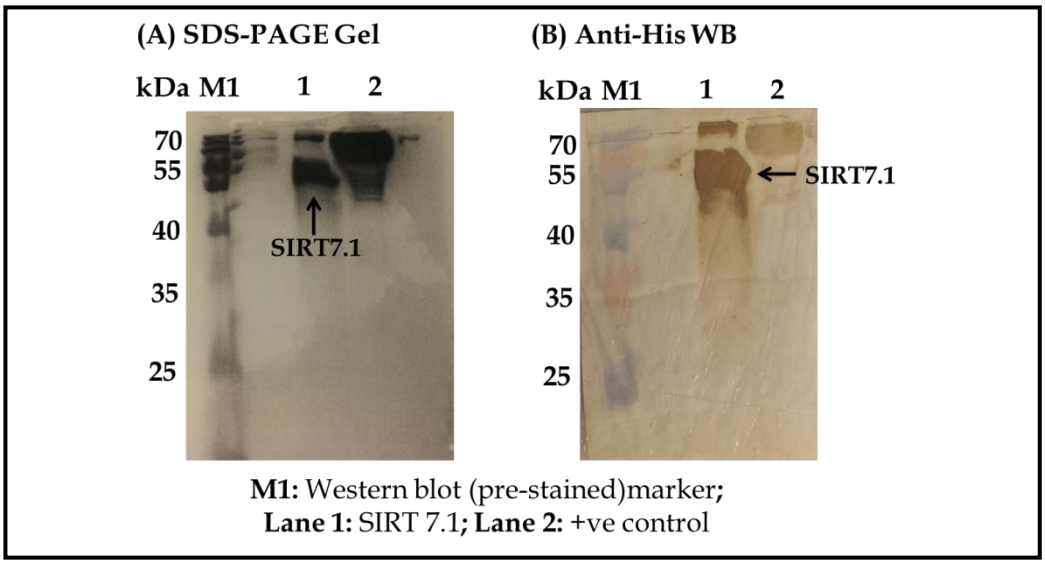


Figure 5.2. 8: Example of SIRT7.1 western blot analysis using anti-His Ab.

Samples of SIRT7.1 analysed by western blotting (WB) using anti-His Ab. The molecular weights of the SDS-PAGE marker (M) and western blot marker (M1) are labelled.

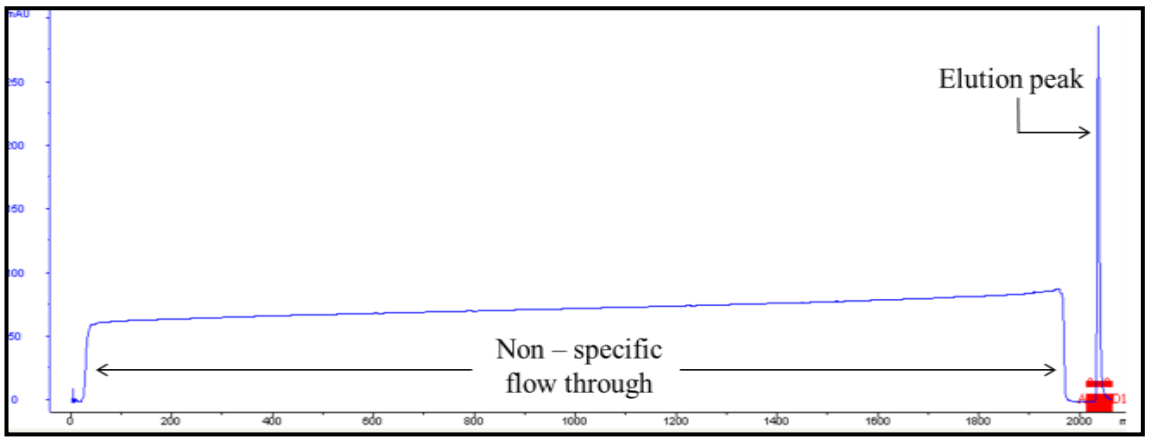


Figure 5.2. 9: SIRT7.1a purification.

Example chromatogram of CIEX by a manually packed SP sepharose column for purification of SIRT7.1a.

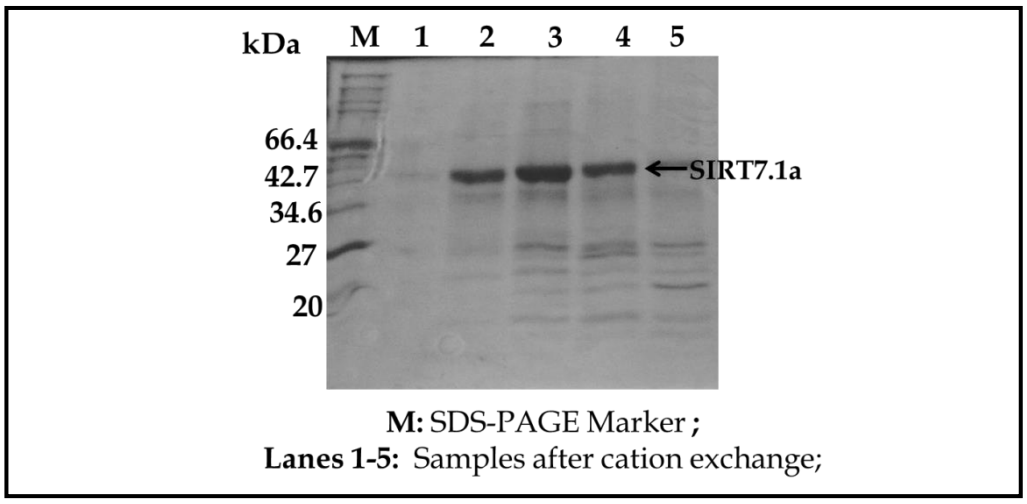


Figure 5.2. 10: Example of SIRT7.1a purified by CIEX analysed by SDS-PAGE.

Purification of SIRT7.1a analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

5.2.3.2. SIRT7 construct SIRT 7.6

The construct SIRT7.6 was designed to enable the soluble expression of SIRT7. The recombinant gene for SIRT7.6 was cloned in a pET41-EK/LIC vector to facilitate the expression of SIRT7 with a covalently linked GST-tag to aid protein expression. The presence of this tag may also facilitate its crystallisation was unsuccessful with His-tagged SIRT7.1.

The primers *SIRT7.6fwd* and *SIRT7.6rev* were used for PCR amplification of the *SIRT7.1* expression vector where a band of ~1400 bp was observed on analysis of the sample on an agarose gel [Figure 5.2.11]. The resulting gene of interest was used for annealing with the vector pET41-Ek/LIC followed by the successful transformation into *E. coli* TOP10 cells and DNA extraction. On verification of the DNA sequencing, transformation of SIRT7.6 recombinant plasmid into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells were carried out. The alignment of the translated DNA sequence of SIRT7.6 with SIRT7 is given in appendix E2.2. Expression trials were initially performed in LB and TB media which showed expression of SIRT7.1 as inclusion bodies [Figure 5.2.11]. Further expression trials in nine additional media at 37 ^{o}C yielded similar results [Figure 5.2.12]. To confirm the insoluble expression observed was not due to expression at high temperature, expression trials at 16 ^{o}C in superbroth media were carried out which showed no expression of SIRT7.6 [Figure 5.2.12].

As soluble expression of SIRT 7.6 was sought, the recombinant plasmid for SIRT7.6 was transformed into *E. coli* ArcticExpress (DE3) RIL cells (referred to as SIRT7.6a) and expression trials in LB media showed soluble expression of the protein [Figure 5.2.13]. *E. coli* ArcticExpress (DE3) RIL cells are engineered to enable efficient expression of heterologous protein at lower temperatures to ensure its soluble expression. However, leaky expression was observed in addition to the chaperonin Cpn60 getting co-purified with SIRT7.6a during initial purification trials. Addition of 10 mM $MgCl_2$ and 10 mM glucose to the LB media during expression solved both the problems [Figure 5.2.13]. Successful purification of SIRT7.6a using GST-AC followed by desalting and SEC was achieved [Figure 5.2.14 and 5.2.15]. An yield of 10-15 mg of the GST-tagged SIRT7.6 was obtained using this method. This purified protein was used to set up crystallisation experiments discussed in section 5.3.2.5.

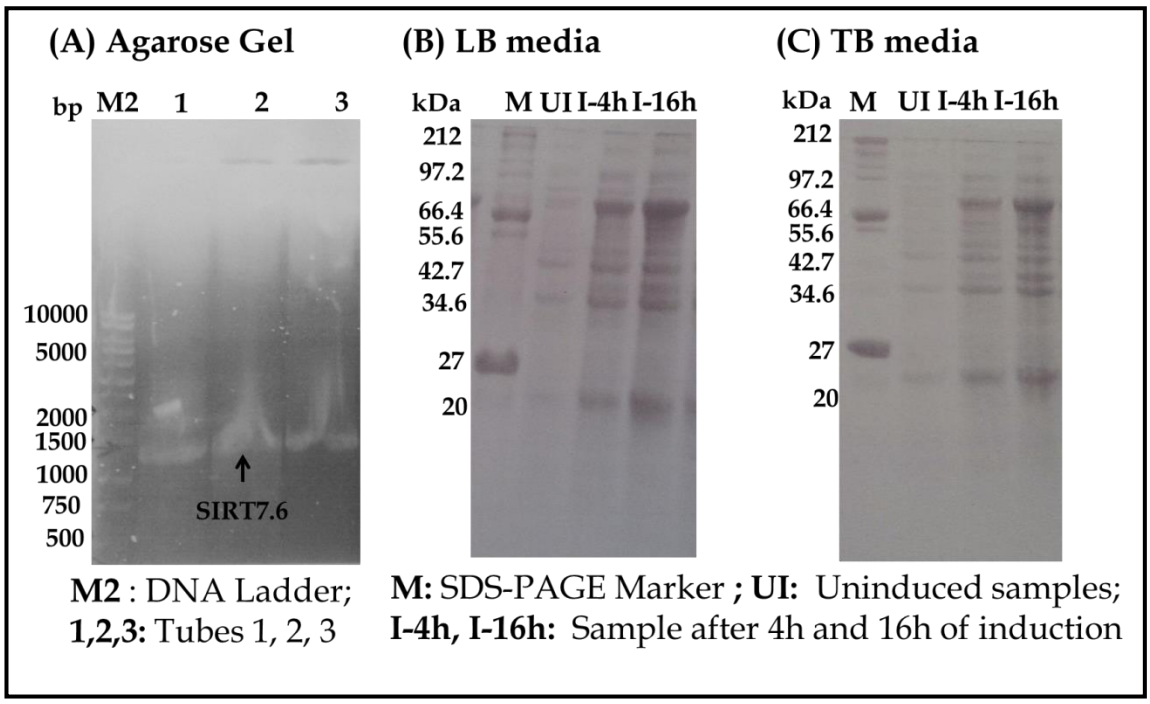


Figure 5.2. 11: Cloning and expression trials for SIRT7.6.

(A) Agarose gel of the PCR product for SIRT7.6. (B) Expression trials for SIRT7.6 in LB and TB at 37 ^{o}C. The base pairs of the DNA ladder (M2) and molecular weights of the SDS-PAGE marker (M) are labelled.

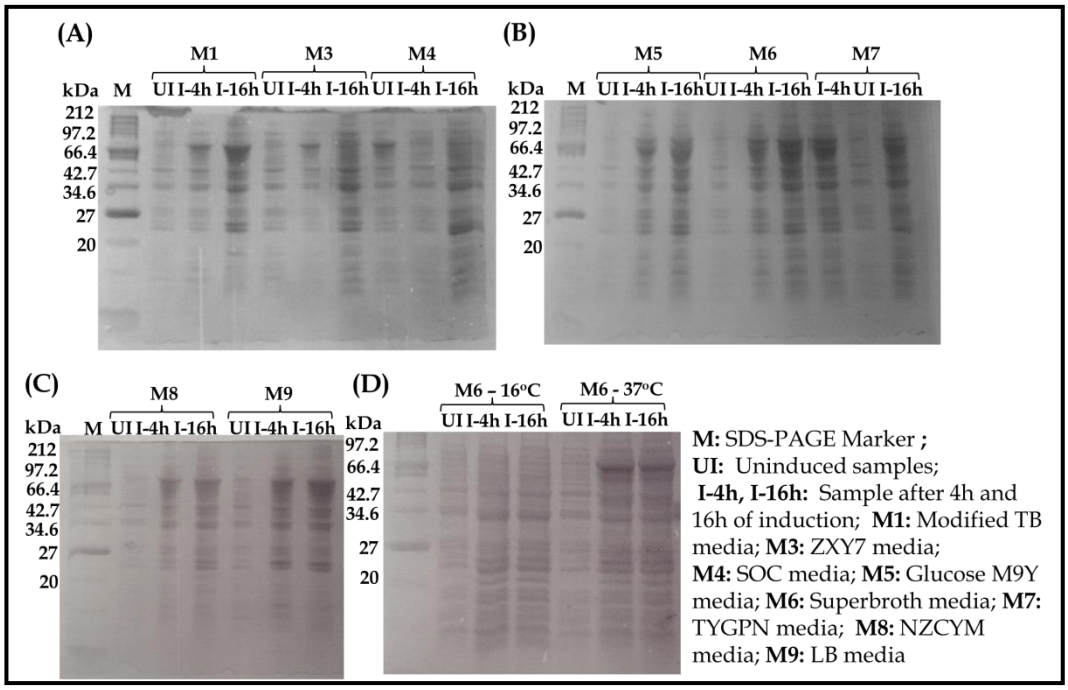


Figure 5.2. 12: Expression trials for SIRT7.6 analysed by SDS-PAGE.

(A-C) Samples from SIRT7.6 expression trials at 37 ^{o}C (D) Large scale expression in superbroth at 16 ^{o}C and 37 ^{o}C. The molecular weights of the SDS-PAGE marker (M) are labelled.

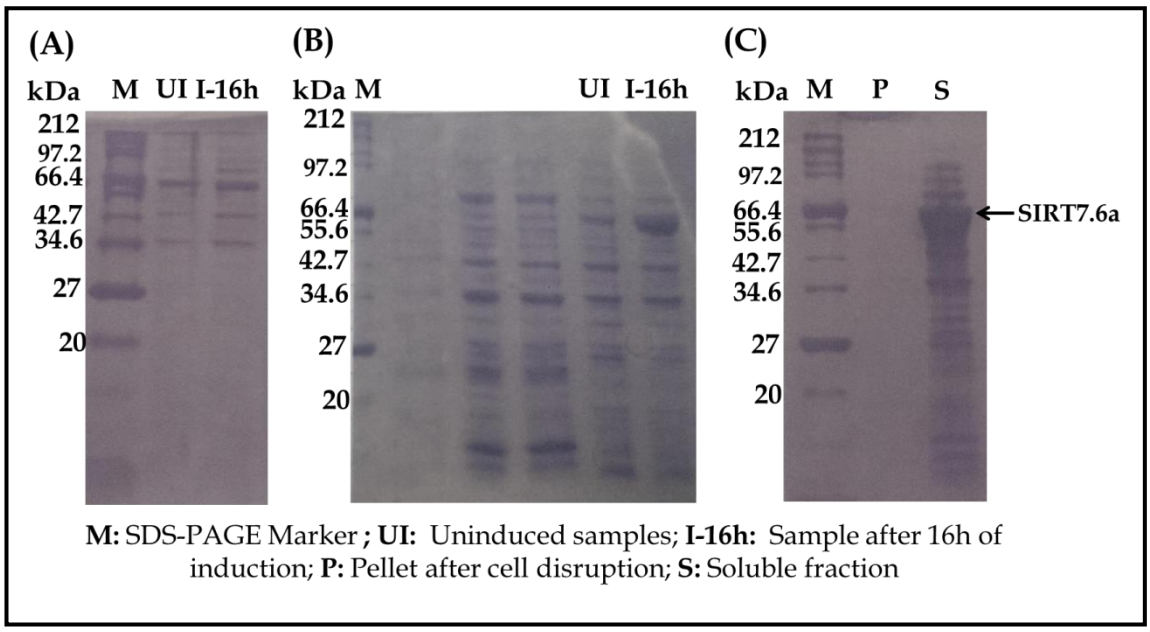


Figure 5.2. 13: Example of SDS-PAGE analysis of expression of SIRT7.6a in LB media.

Cultures after induction incubated at 16 ^{o}C. (A) Expression in LB media. (B) Expression in modified LB media (+ 10 mM $MgCl_2$ and 10 mM glucose). (C) SDS-PAGE of SIRT7.6a after cell disruption.

Anti-GST Ab was used for western blot analysis of SIRT7.6a to confirm the identity of the protein [Figure 5.2.14]. As dialysis of SIRT7.6a in water for intact mass spectrometric analysis resulted in precipitation of the protein, tryptic digest of in-gel protein for mass spectrometric analysis was employed which confirmed the identity of the protein On treatment with enterokinase restriction enzyme to cleave the GST-tag from the purified protein, precipitation of the protein was observed. Thus, all crystallisation experiments were set up using the GST-tagged protein.

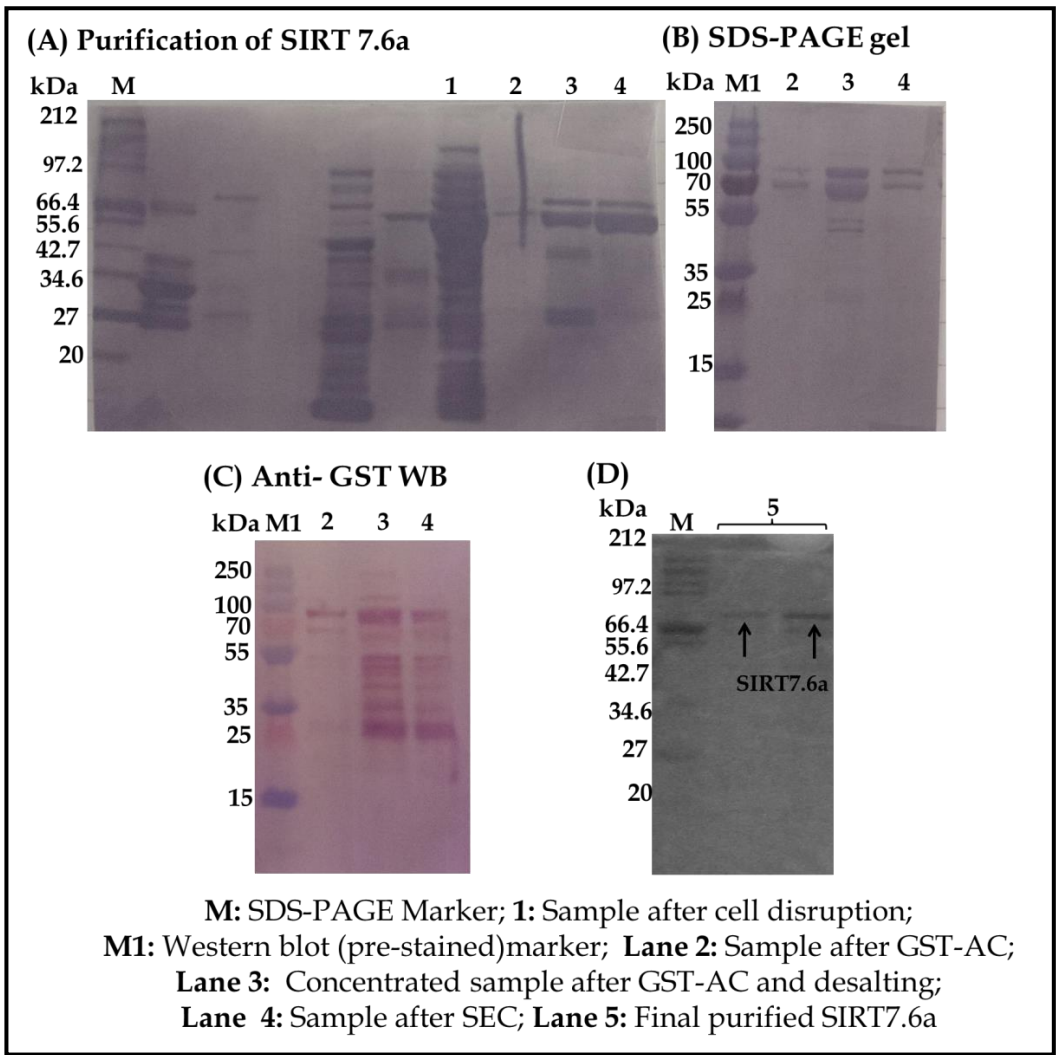


Figure 5.2. 14: Example of purification of SIRT7.6a analysed by SDS-PAGE and western blot analysis using anti-GST Ab.

The molecular weights of the SDS-PAGE marker (M) and the western blot markers (M1) are labelled.

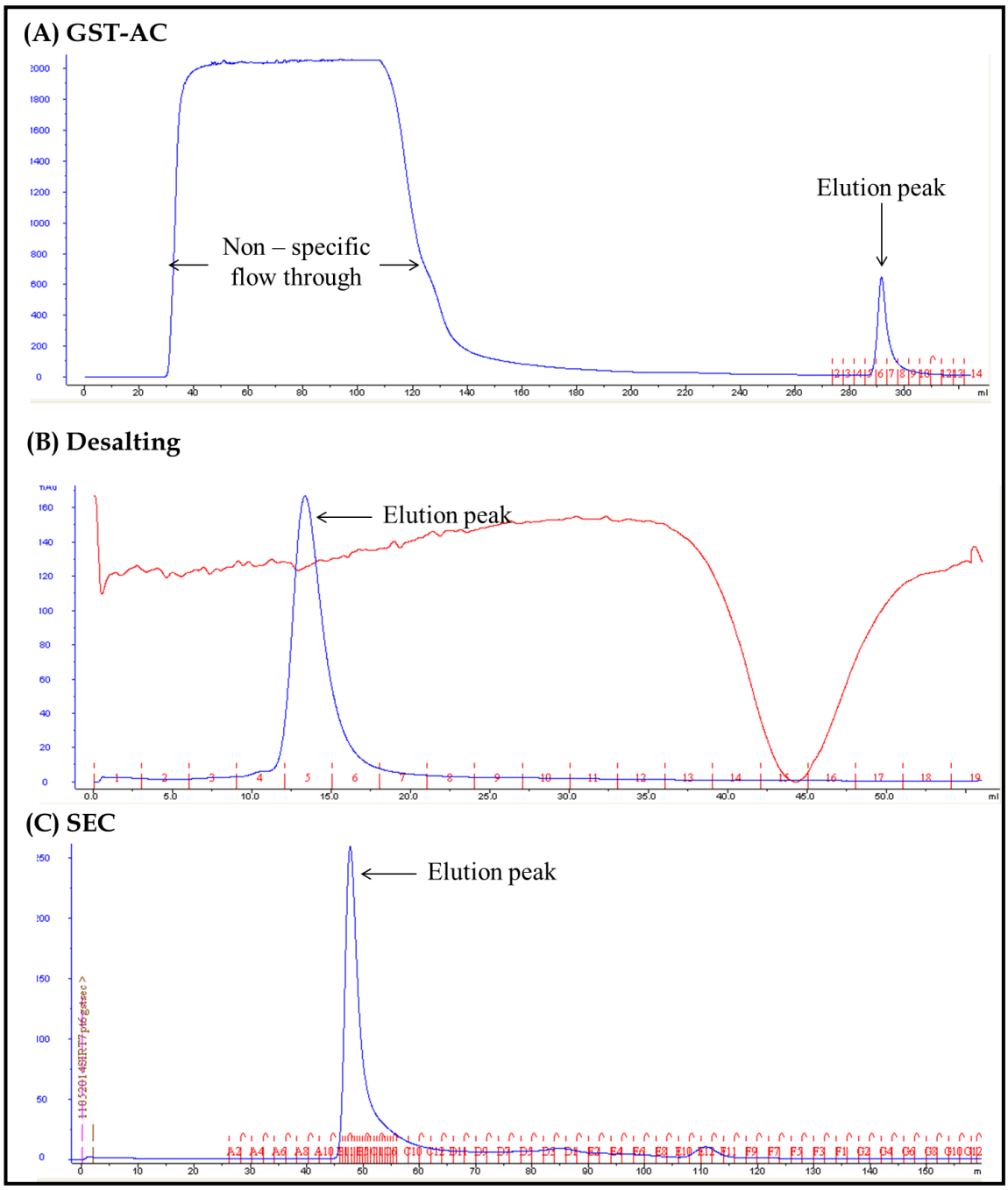


Figure 5.2. 15: Purification of SIRT7.6a.

Example chromatogram of (A) GST-AC utilising the GSTrap FF column for purification of GST-tagged SIRT7.6a followed by (B) desalting and (C) SEC using a superdex 200 column. Blue line represents the absorbance at A_{280} and the red line represents the conductivity.

5.2.3.3. SIRT7 construct SIRT7.7

As GST-tag cleavage of SIRT7.6a, to obtain untagged SIRT7, resulted in precipitation of the protein, SIRT7.7 construct was designed. The recombinant gene for construct SIRT7.7 was cloned in a pGEx-6P-1 vector to facilitate the expression of SIRT7 with a covalently linked GST-tag to aid protein expression. The primers *SIRT7.7fwd* and *SIRT7.7rev* were used for PCR amplification of *SIRT7.1* expression vector where a band of ~1400 was observed on analysis of the sample on an agarose gel [Figure 5.2.16]. The resulting gene of interest was used for annealing with the PCR Blunt vector and was successfully transformed into *E. coli* TOP10 cells for DNA extraction and sequencing. Positive results for forward DNA sequencing was obtained. The alignment of the translated DNA sequence of SIRT7.7 in the PCR blunt vector with SIRT7 is given in appendix E2.3.

On verification of the DNA sequencing, double digestion of pGEx-6P-1 vector and SIRT7.7 PCR blunt vector with EcoRI and XhoI restriction enzymes were carried out [Figure 5.2.16]. A drop out at ~1400 bp for the blunt vector and ~4000 bp for pGEx-6P-1 was observed. These bands were extracted and ligation of the double digested vector with drop out from the blunt vector was carried out. The ligated recombinant plasmid was transformed into Top 10 cells and the DNA was extracted for sequencing. Positive results for forward DNA sequencing was obtained. The alignment of the translated DNA sequence of SIRT7.7 in pGEx-6P-1 vector with SIRT7 is given in appendix E2.4. On verification of the DNA sequencing, the plasmid was transformed into *E. coli* ArcticExpress (DE3) RIL cells to enable soluble expression of the protein at low temperatures. To avoid leaky expression, 10 mM glucose was added to the LB media during expression [Figure 5.2.17]. Purification of SIRT7.7 was attained using GST-AC followed by desalting. [Figure 5.2.17 and 5.2.18]. Tryptic digest of in-gel protein for mass spectrometric analysis confirmed the identity of the protein. PreScission protease treatment for GST-tag cleavage was performed and the resulting untagged protein (~41 kDa) was collected in the flow through. An yield of ~3-5 mg of untagged SIRT7 per litre of cell culture was obtained.

This sample was used for crystallisation experiments discussed in section 5.2.3.5.

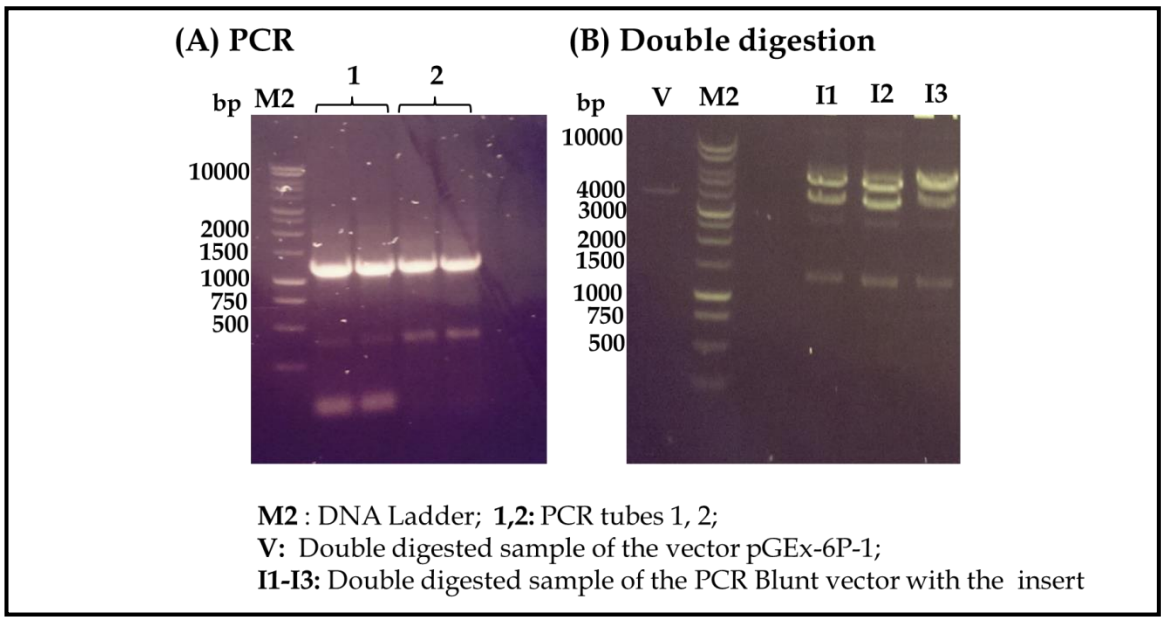


Figure 5.2. 16: SIRT7.7. cloning.

(A) Agarose gel of the PCR product for SIRT7.7. (B) Double digestion for SIRT7.7. The base pairs of the DNA ladder (M2) are labelled.

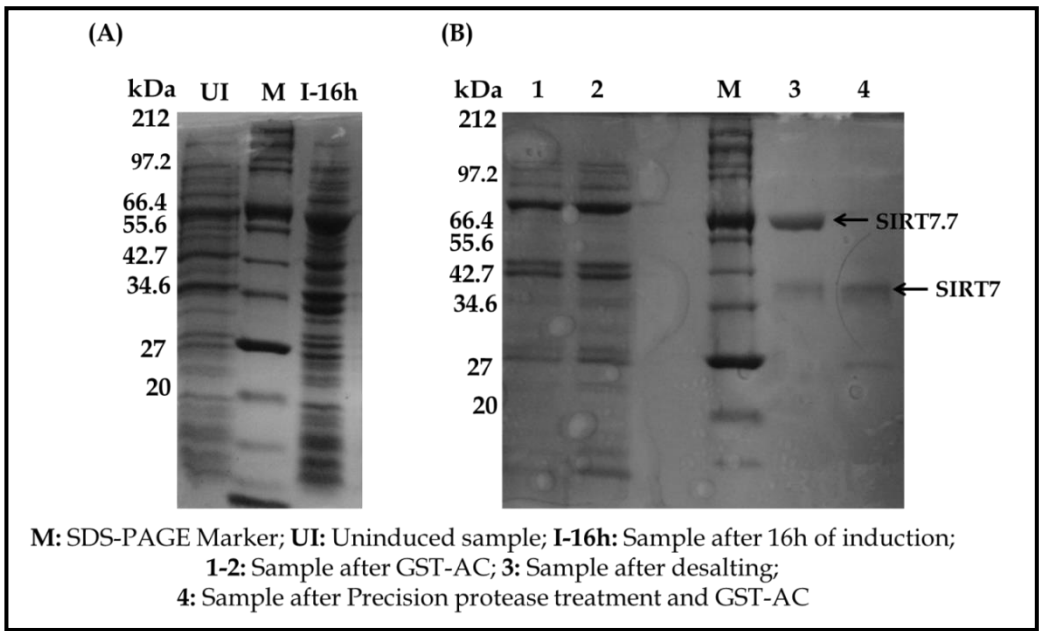


Figure 5.2. 17: Example gel for expression and purification of SIRT7.7 analysed by SDS-PAGE.

(A) SIRT7.7 expression in LB media. (B) Purification of SIRT7.7 by GST-AC, desalting followed by PreScission Protease treatment and GST-AC. The molecular weights of the SDS-PAGE marker (M) are labelled.

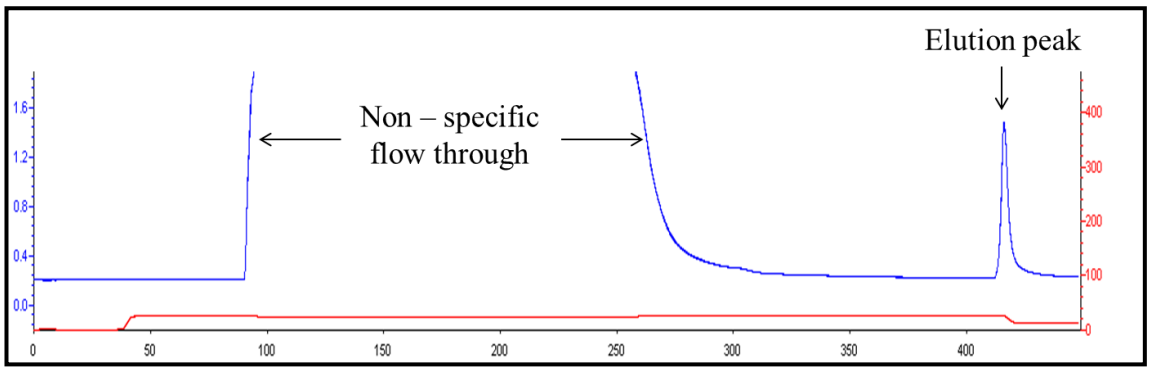


Figure 5.2. 18: SIRT7.7 purification.

Example chromatogram using GSTrap FF column for purification of GST tagged SIRT7.7. Blue line represents the absorbance at A_{280}.

5.2.3.4. SIRT7 activity assay

The SIRT7 activity assay was performed for SIRT7.1a to ensure correct refolding of the protein after denaturation during its purification. Incorrect refolding of the protein and the resulting inactive protein remain two major concerns for crystallisation of proteins expressed as inclusion bodies.

To confirm that the protein was active, an assay was designed to detect the binding of NAD^+ to SIRT7. This was done through western blot analysis using biotinylated NAD^+ which binds to the protein and is detected using HRP-conjugated streptavidin. The DAB colorimetric detection method was used to visualise the final product. Figure 5.2.19 shows the western blot membrane after detection with DAB. The detection of the SIRT7-NAD^+ complex at the expected molecular weight for SIRT7.1a was indicative that the purified protein was active.

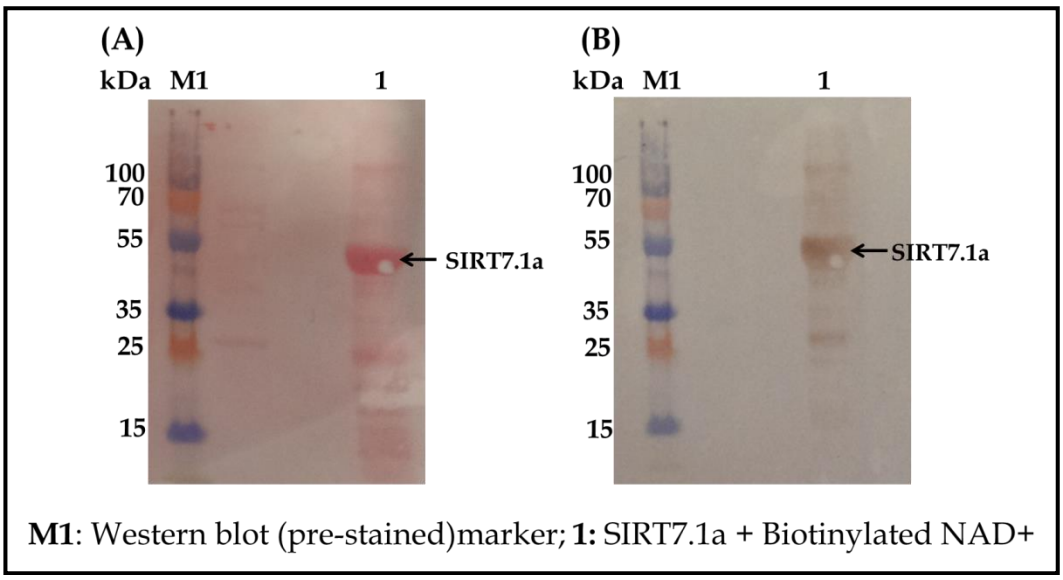


Figure 5.2. 19: NAD+ activity assay for SIRT7.1.

(A) Membrane after transfer stained with Ponceau S stain. (B) Membrane after detection using HRP-conjugated streptavidin followed by colorimetric detector, DAB. The molecular weights of the western blot marker (M1) are labelled.

5.2.3.5. Crystallisation trials for SIRT7

In an attempt to determine the experimental variables and a suitable condition for crystal growth, various experiments based on a coarse grid were designed. Crystallisation trials for SIRT7 started with setting up of sparse-matrix commercial conditions (Section 3.7.2). In an attempt to crystallise SIRT 7, purified protein from constructs SIRT7.1, SIRT7.1a, SIRT 7.6, SIRT7.6a and SIRT7.7 were used for setting up automated preliminary crystallisation screening. All 12 commercial crystallisation screens (12 screens x 96 conditions each x 3 ratios x 3 temperatures x 5 protein constructs x 5 buffers) were set up to facilitate the crystallisation of native SIRT7. The vapour diffusion sitting drop method was the preferred crystallisation technique. A variety of conditions were tried at different temperatures (4 ^{o}C, 16 ^{o}C and 21 ^{o}C), with drops containing the protein solution and the reservoir solution in a ratio of 1:1, 1:2 and 1:3.

Of the various conditions used for crystallisation screening, many deceptive salt crystals were obtained which was confirmed on X-ray exposure at the synchrotron. Some screens including Clear Strategy I HT-96, Clear Strategy II

HT-96 and Morpheus HT-96 produced clear drops while screens such as the Stura Footprint Combination HT-96, JCSG-*plus* HT-96 and Morpheus HT-96 predominantly produced cloudy amorphous precipitate.

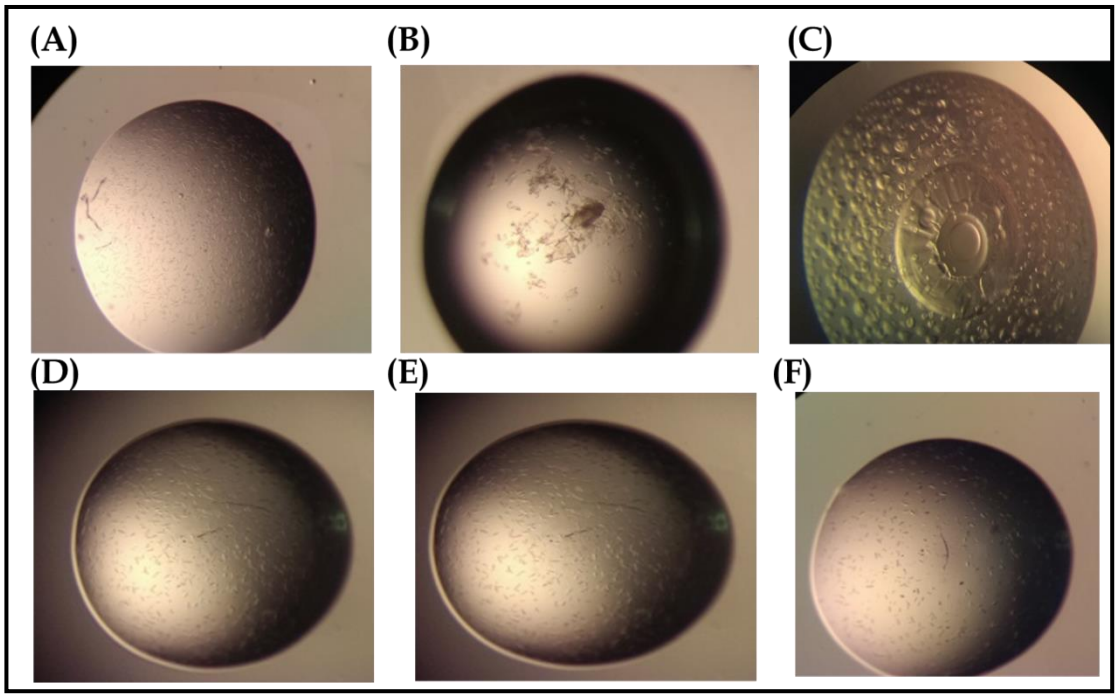


Figure 5.2. 20: Examples of crystallisation hits observed for SIRT7.

(A) Native SIRT7.6a crystals (B) SIRT7.1 in complex with ADPR (C and F) SIRT7.6a in complex with ADPR (D) Native SIRT7.6a crystals (E) Native SIRT7.1a Conditions: (A) B9-Heavy + Light twin pack (B-C) E9-Heavy + Light twin pack (D) C9 JCSG-*plus* HT-96, (E) A9-Heavy + Light twin pack. (F) D9-Heavy + Light twin pack.

However, conditions B9 (0.2 M sodium citrate , 0.1 M Tris pH 8.5, 30% v/v PEG 400) and D9 (0.2 M sodium citrate , 0.1 M Tris pH 8.5, 21% v/v PEG 400) of the Heavy + Light twin pack HT-96 yielded microcrystals for SIRT7.6a at 4 mg/mL after 95 days of incubation at 16 ^{o}C [Figure 5.2.20]. In addition to a systematic screening of these conditions, crystallisation trials for SIRT7.6a in complex with NAD and with ADPR were also designed. Commercial crystallisation screens for SIRT7.6a in complex with NAD and SIRT 7.6a in complex with ADP-ribose were set up at molar ratios of 1:1, 1:2, 1:4 and 1:10 of protein to ligand (12 screens x 96 conditions each x 4 ratios of protein to ligand x 2 protein concentrations x 2 ligands). A protein concentration of 4 mg/mL and 6 mg/mL were opted for these trials. A similar screening with NAD^{+} and ADPR was set

up for SIRT7.1 (12 screens x 96 conditions each x 4 ratios of protein to ligand x 2 protein concentrations x 2 ligands). Of the various screens, very small crystals were obtained for SIRT7.6a in complex with ADPR at a ratio of protein to ligand of 1:4 in A9 (0.2M zinc acetate, 0.1M sodium cacodylate pH 6.5, 18% [w/v] PEG 8000) and D9 (0.2 M sodium citrate , 0.1 M Tris pH 8.5, 21% v/v PEG 400) of the Heavy + Light twin pack HT-96 [Figure 5.2.20]. These crystals took 3 months of incubation at 16 ^{o}C to grow. Similar crystals grew in C9 JCSG-*plus* HT-96 (0.1M Na/K phosphate, 25% v/v 1,2-propanediol, 10% [w/v] glycerol) on incubation at 16 ^{o}C. However, on repeating this screen, it was observed that a protein to reservoir solution ratio of 1:2 was most preferred for crystal growth. These crystals were exposed to X-ray diffraction on beam-line I04 but no diffraction was observed.

Although SIRT7 was expressed and purified, its crystallisation to date has been unsuccessful. Possible reasons for this are discussed in section 5.4.

5.3. Sirtuin 1

5.3.1. Introduction

The human Sir2 orthologue, SIRT1, is an NAD^+ dependent deacetylase that is localised in the nucleus (Imai *et al.*, 2000, Vaziri *et al.*, 2001). It is the most widely studied member of the sirtuin family. SIRT1 interacts with a wide range of substrates and has been identified in several pathological conditions including ageing and age-related diseases. SIRT1 deacetylates a number of histone and non-histone proteins including histone H3 and H4, p53, NF-κB, PGC-1α, HIF1α, among others (Blander and Guarente, 2004, Feldman *et al.*, 2012, Houtkooper *et al.*, 2012, McGuinness *et al.*, 2011). The dual role of SIRT1 as a tumour suppressor and a tumour promoter has been a topic of discussion and has eluded our understanding of its role in cancer. The known SIRT1 substrates and the role of SIRT1 in various pathological conditions are discussed in detail in sections 2.4 and 2.6 respectively.

Despite progress over the last decade, relatively less is known about the regulatory mechanism of SIRT1. SIRT1 is actively promoted or inhibited by endogenous proteins, AROS (active regulator of SIRT1) and DBC1 (deleted in breast cancer 1) (Kim *et al.*, 2007, Kim *et al.*, 2008, Zhao *et al.*, 2008).

Table 5.3. 1: Overview of the results of a sequence comparison of SIRT1 with other human sirtuins.

ClustalO was used for sequence comparison (UniProt, 2014).

		Identity (%)	Identical positions
SIRT1	SIRT2	17.806	138
SIRT1	SIRT3	17.25	133
SIRT1	SIRT4	10.013	78
SIRT1	SIRT5	10.78	83
SIRT1	SIRT6	10.39	80
SIRT1	SIRT7	10.35	82

Two splice variants of SIRT1 have been reported, isoform-1 and -2 with 747 and 561 residues, respectively (UniProt, 2014). The canonical SIRT1, isoform 1, is the longest member of the sirtuin family comprised of 747 residues. It is an 81.7

kDa protein with regions for two nuclear localisation signals and two nuclear export signals [Figure 5.3.1]. The deacetylase domain of SIRT1 is the most conserved domain among sirtuins. SIRT2 and SIRT3 are the closest homologues of SIRT1 while SIRT4 and SIRT7 are the least similar [Table 5.3.1].

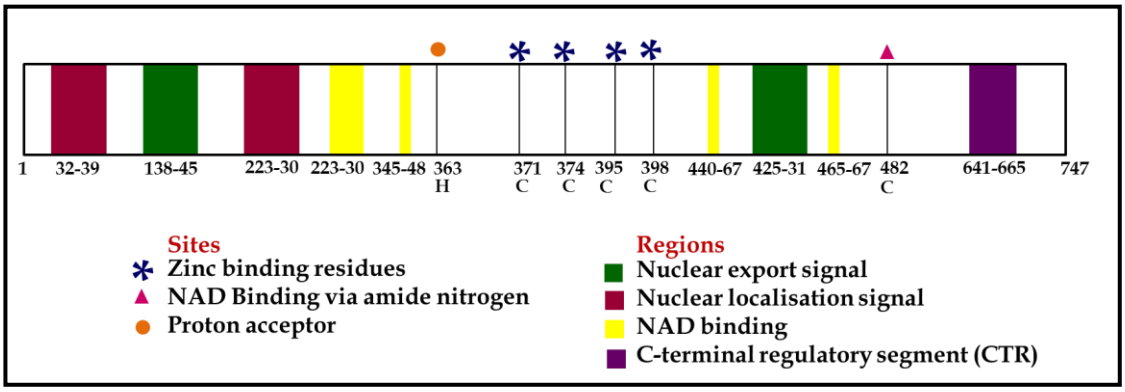


Figure 5.3. 1: Key features of the structure of SIRT1.

A schematic representation of the structure of SIRT1. The structurally important cysteine (C) residues involved in zinc binding are indicated along with the histidine (H) residue that behaves as a proton acceptor.

Of the seven human sirtuins, SIRT1 contains the most extended N- and C-terminal segments. The region between residues 200–500 forms the highly conserved catalytic core domain. However, various regions in the long and mostly unstructured N- and C-termini that flank the SIRT1 catalytic domain have been shown to affect SIRT1 deacetylation activity through autoregulation (Kang *et al.*, 2011, Pan *et al.*, 2012). The terminal segments were shown to potentiate the catalytic efficiency of the core domain by 12-45 fold, with the N-terminal domain contributing significantly to the rate of catalysis and the C-terminal contributing to the enzyme's affinity to NAD^+ (Pan *et al.*, 2012).

In the recent years, three crystal structures of truncated SIRT1 domains have been elucidated. The first crystal structure of SIRT1 determined was the catalytic core domain (residues 214-516) in complex with a NAD^+ analogue and the SIRT1 inhibitor EX-527 analogue (Zhao *et al.*, 2013). This was followed by the publication of two additional structures of SIRT1 (Davenport *et al.*, 2014). The first showed residues 234-510 and additional residues 641–665 forming the

C-terminal regulatory segment while the second structure was a high resolution structure of the apo-form of SIRT1. Davenport *et al*. suggested that the C-terminal regulatory segment (residues 641-665) contributes to SIRT1 activity by influencing the conformational changes that are required for substrate and cofactor binding.

With the aim to elucidate the crystal structure of SIRT1 in its native form and in complex with other ligands, structural biology work on SIRT1 (isoform-1) was initiated. The following section provides details of the work done to this end.

5.3.2. Methods

5.3.2.1. Cloning, Expression and Purification

Table 5.3.2 enlists the constructs of SIRT1 that were designed for its expression with covalently linked affinity tags to aid expression and purification of the protein. The vector maps are provided in appendix A. The final optimised cloning, expression and protein purification protocols for each construct are given below.

Table 5.3. 2: Summary of SIRT1 constructs

Construct	Protein insert (residues)	Vector	Tag	Mol. Wt. (kDa)
1.1	140-747	pTriEx-4 Ek/LIC	N-term His and S-tag	73.2
1.5	140-747	pET-41 Ek/LIC	N-term GST, His and S-tag	99.3
1.6	160-747	pET-41 Ek/LIC	N-term GST, His and S-tag	97.0
1.7	160-747	pTriEx-4 Ek/LIC	N-term His and S-tag	88.4
1.8	160-665	pET-41 Ek/LIC	N-term GST, His and S-tag	70.8
1.9	160-665	pTriEx-4 Ek/LIC	N-term His and S-tag	62.2

SIRT1.1

The gene coding for SIRT1 was cloned into pTriEx-4 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal poly-histidine and S-tag. This recombinant plasmid was named SIRT1.1. The plasmid was transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells. The glycerol stocks of this transformant were available in the laboratory. Cells were grown in AI media with ampicillin at 37 °C for 6 h and then incubated at 16 ^{o}C for 16 h. Cells were harvested after 22 h of inoculation by centrifugation (4 °C at 12,000 g, 10 min). Cell pellets were resuspended in wash buffer (50 mM HEPES pH 7.4, 200 mM NaCl, 15 mM imidazole) and lysed in an One shot model (Constant systems, U.K.). The preparation was then clarified by centrifugation (4 °C at

75,000 g, 25 min). The supernatant was loaded on a 5 mL HisTrap FF column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with wash buffer. The column was washed extensively and a gradient elution of 0-100% over 50 mL (50 mM HEPES pH 7.4, 200 mM NaCl, 500 mM imidazole) was carried out. The eluted fraction was dialysed overnight in wash buffer at 4 °C and loaded on a gel filtration column (16/600 Superdex 200 pg; GE Healthcare, Uppsala, Sweden) pre-equilibrated with final buffer (50 mM HEPES pH 7.4, 200 mM NaCl). The eluted samples were analysed by SDS-PAGE. Anti-SIRT1 Ab and anti-His Ab (Sigma-Aldrich Co., Kent, U.K.) were used to confirm the identity of the protein by western blot analysis.

SIRT1.1a

On confirming that the protein purified was not SIRT1, the recombinant plasmid for SIRT1.1 was transformed into *E. coli* TOP10 cells and the DNA was extracted for sequencing. This transformant was named SIRT1.1a. On verification of DNA sequencing, transformation into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells followed by preparation of fresh glycerol stocks was carried out as described in Section 3.4.2. Extracted DNA was sent for sequencing to Eurofins MWG Operon, Ebersberg, Germany using T7 primer for forward sequencing and TriExDOWN primer for reverse sequencing.

Expression trials were carried out by inoculating AI, LB and TB media with ampicillin and chloramphenicol as described in Section 3.5.1 followed by induction with 1 mM IPTG. Samples were collected at regular intervals and analysed for protein expression by SDS-PAGE.

SIRT1.5

The gene coding for SIRT1 was cloned into pET-41 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal GST-tag. The recombinant plasmid was named SIRT1.5.

The recombinant plasmid *SIRT1.1* was used for PCR amplification as described in Section 3.4.2. The primers used for PCR amplification are given in table 5.3.3, while the components of the reaction mixtures and the parameters for thermal cycling are given in tables 5.3.4 and 5.3.5, respectively.

Table 5.3. 3: Primers used for PCR reaction of SIRT1.5

Primer	**Sequence (5′-3′)**
SIRT1.5fwd	GAC GAC GAC AAG ATG ATT GGG TAC CGA GAT AAC CTT CTG
SIRT1.5rev	GAG GAG AAG CCC GGT TAT GAT TTG TTT GAT GGA TAG TTC ATG TCT GTT ACT

Table 5.3. 4: PCR reaction for SIRT1.5

Reaction	**Tube 1 (µL)**	**Tube 2 (µL)**	**Tube 3 (µL)**
DNA template (SIRT1.1) (100ng/µL)	2	2	2
Primer (*SIRT1.5fwd*) (10 µM)	1.5	1.5	1.5
Primer (*SIRT1.5rev)* (10 µM)	1.5	1.5	1.5
KOD polymerase (1U/µL)	1	1	1
10x Buffer	5	5	5
dNTPs (2mM each)	5	5	5
$MgSO_4$ (25 mM)	3	3	3
Betaine	-	10	-
DMSO	-	-	2
Water (q.s. 50 µL)	31	21	29

Table 5.3. 5: PCR thermal cycle for SIRT1.5

Cycle	**Conditions**
Initial denaturation	94 °C, 2 min
Denaturation	94 °C, 30 sec
Annealing	58 °C, 1 min
Extension	72 °C, 1 min
Back to step 2 for 35x	
Final extension	72 °C, 10 min
Hold at 4°C	

The PCR product was analysed by running the samples on an agarose gel as described in section 3.4.1. and the final clean-up product was cloned into a pET-41 Ek/LIC vector (Novagen, Darmstadt, Germany) with a cleavable N-terminal GST, poly-histidine and S-tag. The final recombinant plasmid was transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells followed by preparation of glycerol stocks as described in Section 3.4.2. Extracted DNA was sent for sequencing to Eurofins MWG Operon, Ebersberg, Germany using T7 primer and S-tag primer for forward sequencing.

Expression trials were carried out by inoculating AI, LB and TB media with kanamycin and chloramphenicol as described in Section 3.5.1 followed by induction with 1 mM IPTG. Samples were collected at regular intervals and analysed for protein expression by SDS-PAGE.

SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9

Recombinant plasmids with further truncation of SIRT1 were prepared to facilitate expression of the truncated protein with either a GST-tag or His-tag. Table 5.3.2 gives an overview of the truncation and vectors used for these constructs.

The recombinant plasmids were prepared as described for SIRT1.5. The primers used for PCR amplification are given in table 5.3.3 and 5.3.6, while the components of the reaction mixtures and the parameters for thermal cycling are given in tables 5.3.7 and 5.3.8, respectively.

Table 5.3. 6: Primers used for PCR reaction of SIRT1.6-1.9.

Primer	Sequence (5′-3′)
SIRT1.6fwd	GAC GAC GAC AAG ATG TGT GAA AGT GAT G
SIRT1.8rev	GAG GAG AAG CCC GGT TAG ACG TCA TCT TC

Table 5.3. 7: PCR reaction of SIRT1.6-1.9.

Reaction	Tube 1 (µL) (160-747)	Tube 2 (µL) (160-747)	Tube 3 (µL) (160-665)	Tube 4 (µL) (160-665)
DNA template (SIRT1.1) (100ng/µL)	2	2	2	2
Primer (Forward) (10µM)	1 *(SIRT1.6fwd)*	1 *(SIRT1.6fwd)*	1 *(SIRT1.6fwd)*	1 *(SIRT1.6fwd)*
Primer (Reverse) (10 µM)	1 *(SIRT1.5rev)*	1 *(SIRT1.5rev)*	1 *(SIRT1.8rev)*	1 *(SIRT1.8rev)*
KOD polymerase (1 U/µL)	1	1	1	1
10x Buffer	5	5	5	5
dNTPs (2mM each)	5	5	5	5
$MgSO_4$ (25 mM)	3	3	3	3
DMSO	-	2	-	2
Water (q.s. 50 µL)	31	29	31	29

Table 5.3. 8: PCR thermal cycle for SIRT1.6-1.9.

Cycle	Temperature (°C)	Time (min)
Initial denaturation	95	2
Denaturation	95	0.5
Annealing	57	1
Extension	72	1
Back to step 2 (Denaturation) for 35x		
Final extension	72	10
Hold at 4°C		

The PCR products were analysed by running the samples on an agarose gel as described in section 3.4.1. The final purified product was ligated with either pTriEx-4 Ek/LIC or pET-41 Ek/LIC vectors (Novagen, Darmstadt, Germany) as indicated in table 5.3.2. The recombinant vectors were transformed into *E. coli* TOP10 cells and the DNA were extracted as described in section 3.4.2. Extracted DNA was sent for sequencing to Eurofins MWG Operon, Ebersberg, Germany using T7 primer and S-tag primer for sequencing of insert in pET-41 Ek/LIC or

T7 primer and TriExDOWN primer for sequencing of insert in pTriEx-4 Ek/LIC. The verified recombinant plasmids were transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells followed by preparation of glycerol stocks as described in Section 3.4.2.

Expression trials were carried out by inoculating AI, LB and TB media as described in Section 3.5.1. Antibiotics, kanamycin and ampicillin, were used for constructs cloned in pET-41 Ek/LIC and pTriEx-4 Ek/LIC vectors, respectively. Samples were collected at regular intervals and analysed for protein expression by SDS-PAGE.

5.3.3. Results

5.3.3.1. SIRT1 construct SIRT1.1

SIRT1.1 is a truncated protein of residues 140-747. The recombinant gene was cloned into pTriEx-4 Ek/LIC vector in order to facilitate the expression with a covalently linked His-tag. The expressed protein is expected to have a theoretical mass of 73.2 kDa.

Initial expression in TB media followed by IMAC and SEC of SIRT1.1 resulted in a protein at ~55 kDa and ~66kDa being observed on a SDS-PAGE [Figure 5.3.2 and 5.3.3]. Western blot analysis using anti-SIRT1 Ab and anti-His Ab confirmed that the purified protein was not SIRT1.

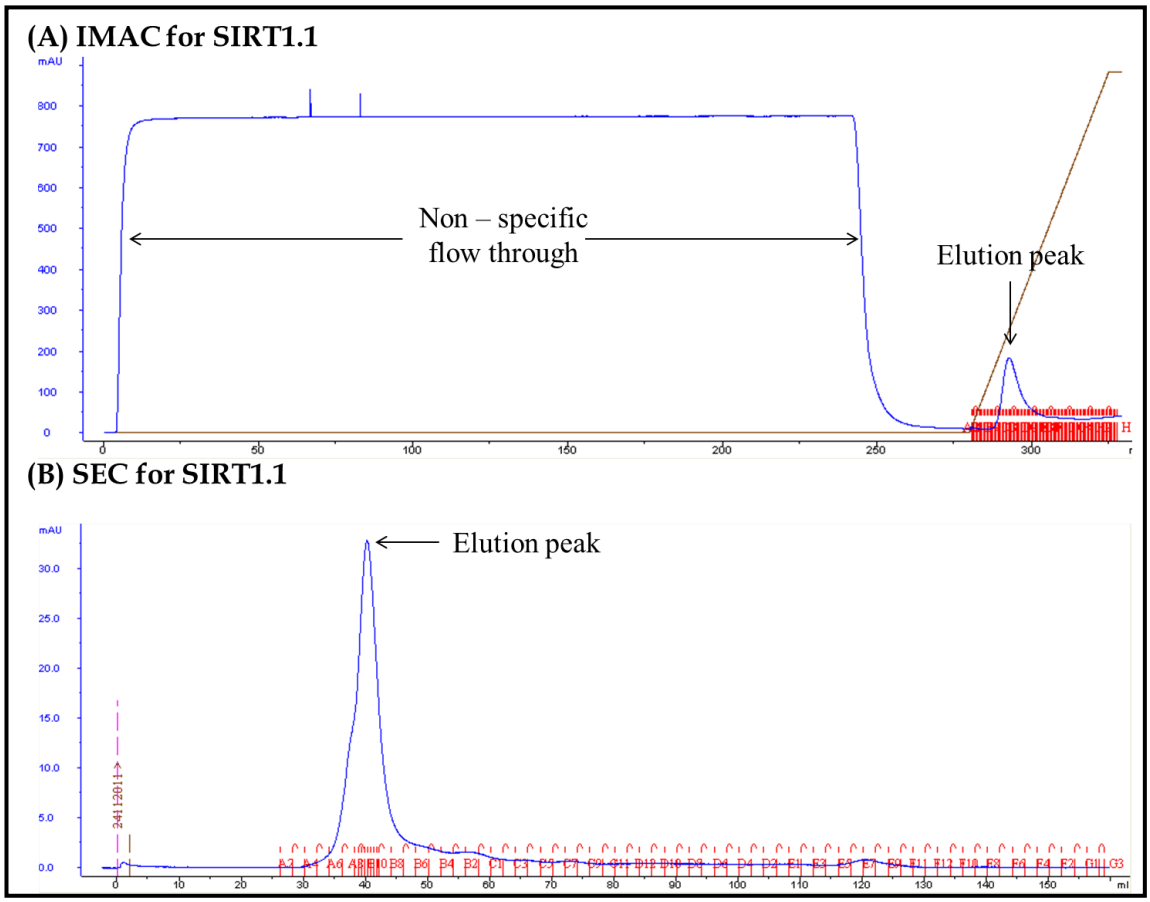


Figure 5.3. 2: SIRT1.1 purification.

Example chromatogram of (A) IMAC utilising the HisTrap FF column for purification of His-tagged SIRT1.1 followed by (B) SEC using a superdex 200 column.

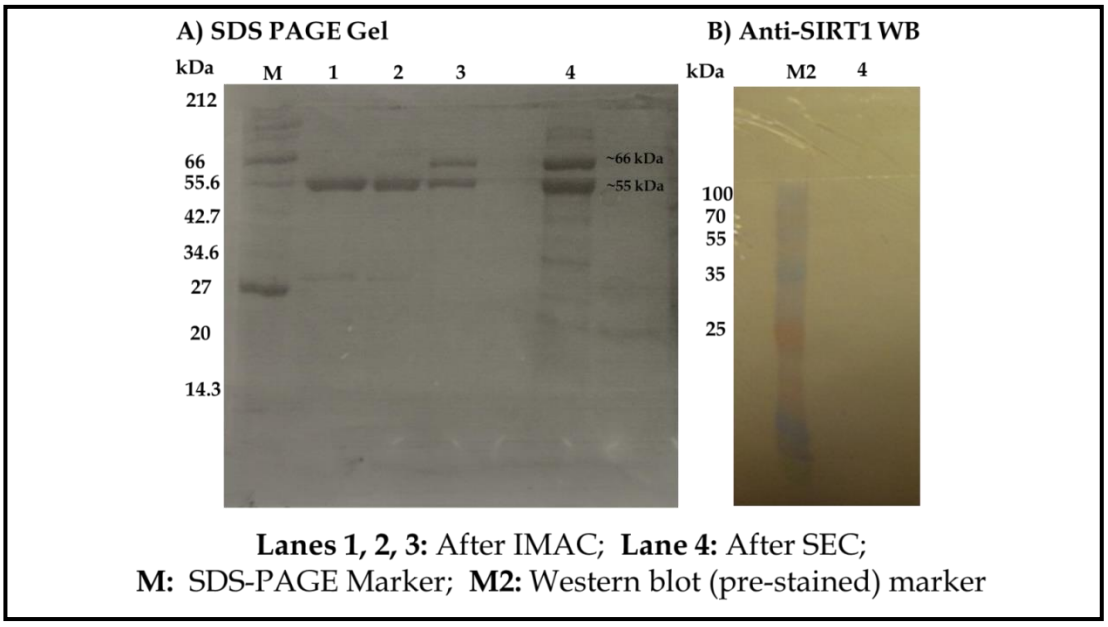


Figure 5.3. 3: SIRT1.1 purification and western blot analysis.

(A) Purification of SIRT1.1 after IMAC and SEC analysed by SDS-PAGE. (B) Western blot analysis of SIRT1.1 after SEC. The molecular weights of the SDS-PAGE marker (M) and western blot marker (M2) are labelled.

DNA extracted from the *E. coli* TOP10 cells glycerol stock was used for successful DNA sequencing (referred to as SIRT1.1a). Positive results for forward and reverse DNA sequencing were obtained. The alignment of the translated DNA sequences of SIRT1.1 with SIRT1 are given in appendix E3.1 and E3.2. The recombinant plasmid was successfully transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells and expression trials were performed in LB, TB and AI media at 37 ^{o}C. On analysis by SDS-PAGE no expression of SIRT1 at 73.2kDa was observed [Figure 5.3.4].

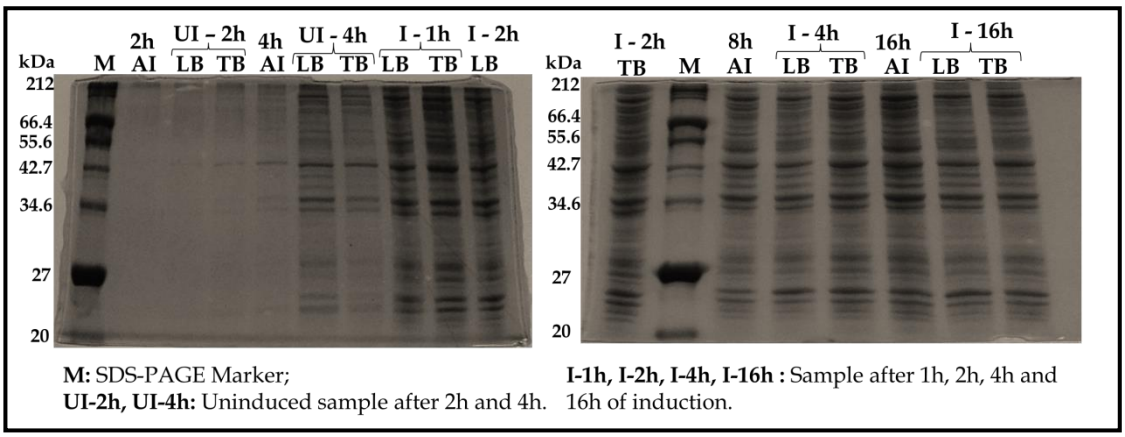


Figure 5.3. 4: Expression trials for SIRT1.1a analysed by SDS-PAGE.

SIRT1.1a expression in LB, TB and AI media at 37 ^{o}C analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

5.3.3.2. SIRT1 construct SIRT1.5

Following SIRT1.1a, the recombinant gene for construct SIRT1.5 was cloned in a pET41-EK/LIC vector to facilitate the expression of SIRT1 with a covalently linked GST-tag to aid soluble protein expression. The primers *SIRT1.5fwd* and *SIRT1.5rev* were used for PCR amplification of *SIRT1.1* expression vector where a band of ~1800 bp was observed on analysis of the sample on an agarose gel [Figure 5.3.5]. The resulting gene of interest was used annealed with pET41-Ek/LIC vector followed by the successful transformation into *E. coli* TOP10 cells and DNA extraction. On verification of the DNA sequencing, transformation of SIRT1.5 recombinant plasmid into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells were carried out. The alignment of the translated DNA sequences of SIRT1.5 with SIRT1 is given in appendix E3.3. Expression trials were performed in LB, TB and AI media at 16 ^{o}C and 37 ^{o}C which showed no expression of SIRT1 at 99.3kDa on analysis by SDS-PAGE [Figure 5.3.5 and 5.3.6].

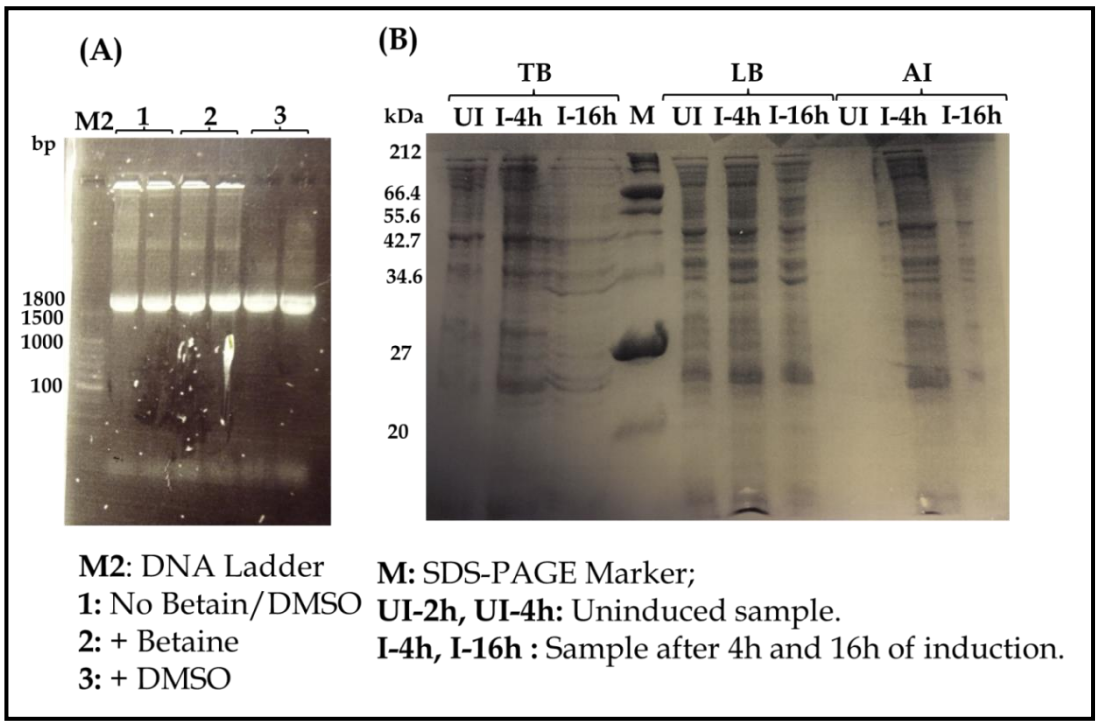


Figure 5.3. 5: Cloning and expression trials for SIRT1.5.

(A) Agarose gel of the PCR product for SIRT1.5. (B) Expression trials for SIRT1.5 in LB, TB and AI media at 37 ^{o}C. The base pairs of the DNA marker (M2) and molecular weights of the SDS-PAGE marker (M) are labelled.

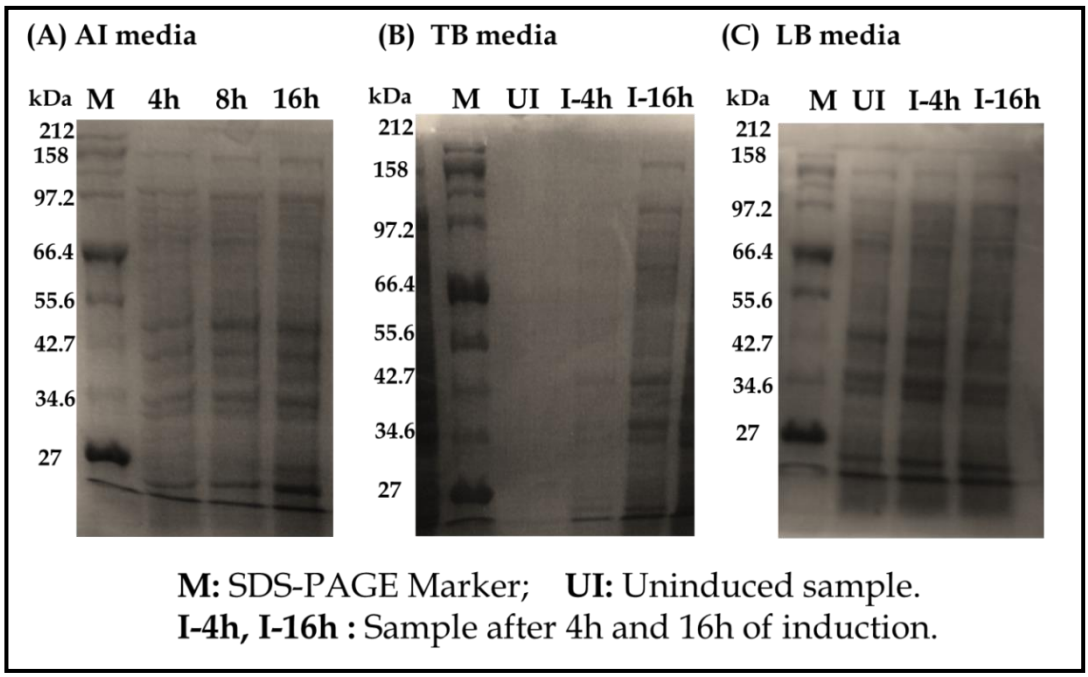


Figure 5.3. 6: Expression trials for SIRT1.5.

Expression trials for SIRT1.5 in LB, TB and AI media at 16 ^{o}C analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

5.3.3.3. SIRT1 constructs SIRT1.6, SIRT1.7, SIRT 1.8 and SIRT1.9

As no expression of SIRT1 (residues 140-747) was observed, constructs with further truncation of SIRT1 to exclude the predicted unstructured regions including the complete nuclear export signal at the N-terminal end and with or without the residues following the C-terminal regulatory segment [Table 5.3.2] were designed.

The primers *SIRT1.6fwd* and *SIRT1.5rev* for residues 160-747 and primers *SIRT1.6fwd* and *SIRT1.8rev* for residues 160-665 were used for successful PCR amplification of *SIRT1.1* expression plasmid [Figure 5.3.7] and final PCR product after clean-up was used for annealing with the respective expression vectors as indicated in table 5.3.2 followed by the successful transformation into *E. coli* TOP10 cells for DNA extraction. The extracted DNA was sent for sequencing analysis. On verification of the sequencing results, SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 recombinant plasmids were transformed into *E. coli* Rosetta-gami™ 2(DE3)pLysS cells. Positive results for forward and reverse DNA sequences were obtained. The alignment of the translated DNA sequences of SIRT1.6-SIRT1.9 with SIRT1 is given in appendix E3.4-E3.9. Expression trials were performed in LB, TB and AI media at 16 ^{o}C and 37 ^{o}C which showed no expression of SIRT1 at the expected molecular weights on analysis by SDS-PAGE [Figure 5.3.8 – 5.3.10].

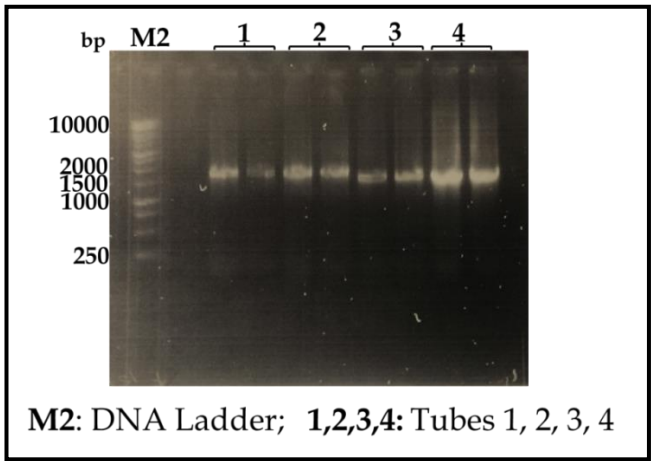


Figure 5.3. 7: Cloning of SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9.

PCR products for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 analysed on an agarose gel. The base pairs of the DNA ladder (M2) are labelled.

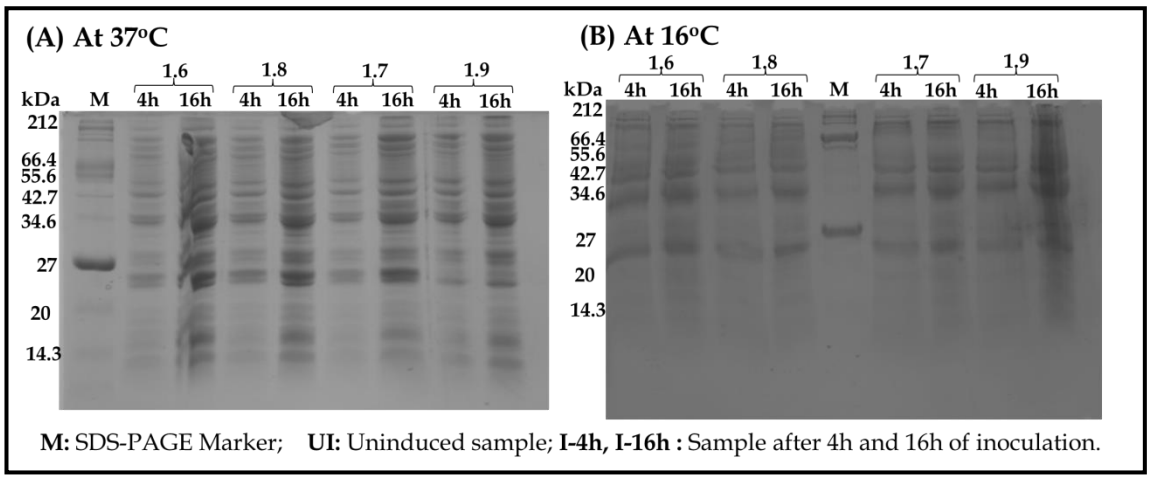


Figure 5.3. 8: Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in AI media.

Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in AI media analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

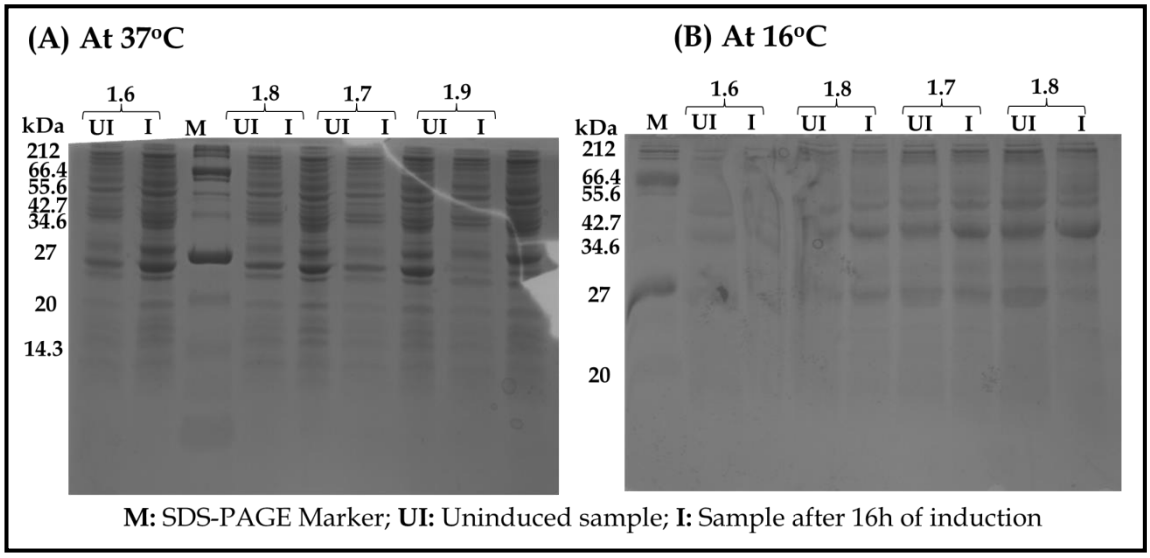


Figure 5.3. 9: Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in LB media.

Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in LB media analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

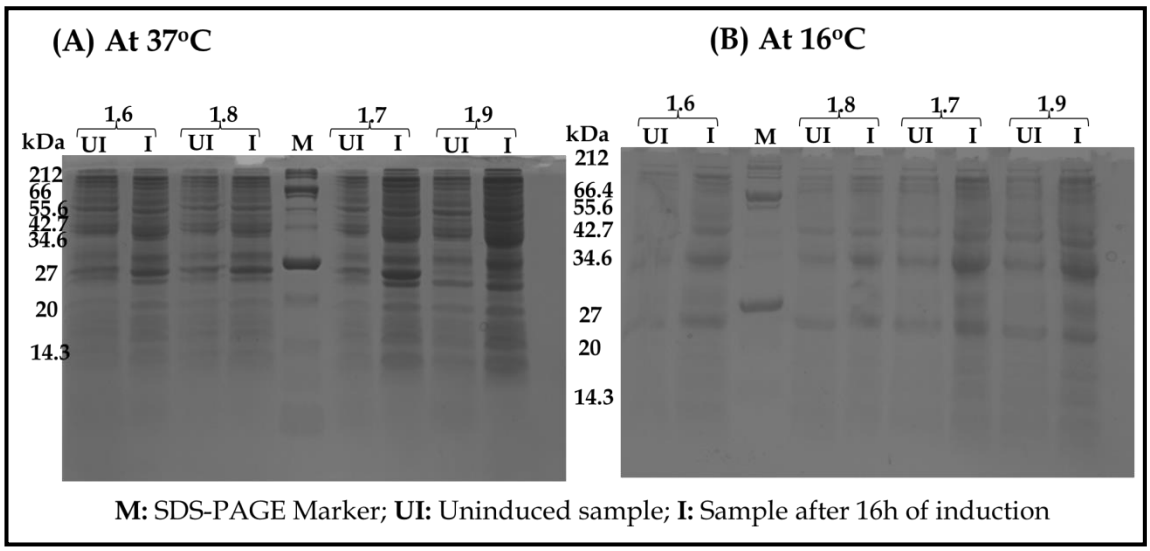


Figure 5.3. 10: Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in TB media.

Expression trials for SIRT1.6, SIRT1.7, SIRT1.8 and SIRT1.9 in TB media analysed by SDS-PAGE. The molecular weights of the SDS-PAGE marker (M) are labelled.

To date, expression and purification of SIRT1 has been unsuccessful. Possible reasons for this are discussed in section 5.4.

5.4. Discussion

Crystallisation of proteins to elucidate their molecular structure using X-ray crystallography is accomplished via a sequential multi-stage process. One of the largest challenges in obtaining crystal structures is the need for milligram quantities of highly purified, soluble, active protein. Once this hurdle is overcome, the next obstacle is the crystallisation of the protein in a form that yields diffraction quality crystals. This study demonstrates the challenges faced in the structural studies on sirtuins and the methods employed to overcome these barriers.

The members of the sirtuin family have been extensively studied in order to understand their role in NAD^+ dependent acylation of diverse substrates and their implication in ageing and ageing related diseases (as discussed in section 2.5). Several sirtuin proteins were cloned, expressed and purified in an attempt to crystallise the proteins in their native forms as well as in complex with various substrates and inhibitors.

5.4.1. Protein expression and purification

Overexpression of eukaryotic proteins in bacteria in their soluble form has been a major bottleneck in structural biology (Smyth *et al.*, 2003). Over the years various large affinity tags such as SUMO (11 kDa), GST (26 kDa), and MBP (40 kDa) among others, have been used to overcome this problem. However, the subsequent cleavage of these tags may lead to problems such as low yield, precipitation of the target protein, and/or failure to recover active or structurally intact protein (Baneyx, 1999). In addition, protein cleavage often requires tedious optimisation and the proteases can be very costly. The alternative, to circumvent the cleavage and re-purification steps, is to leave the affinity tag in place for crystallisation. Unfortunately this brings about a new challenge as multidomain proteins are less conducive to forming well-ordered, diffracting crystals, which is likely to be due to the conformational heterogeneity allowed by flexible linker regions (Bucher *et al.*, 2002). Thus, smaller affinity tags, such as the His-tag, are often selected for structural

biology applications as they do not significantly increase the size of the protein and cleavage of the tags is not essential for crystal growth. However, the usage of smaller tags can compromise soluble overexpression of the protein. Each type of tag has its own advantages and disadvantages (Smyth *et al.*, 2003). Hence, the careful selection of a suitable affinity tag to aid soluble expression, purification, tag cleavage (if required) and crystallisation is a crucial step in structural biology. With these parameters in mind, various sirtuin constructs were designed which included truncated forms of the protein, as well as a variety of affinity tags and *E. coli* expression cell lines.

SIRT2

Four constructs with a range of SIRT2 truncations and affinity tags were expressed and purified for crystallisation trials [Table 5.1.2]. Mass spectrometric analysis and western blot analysis, using anti-SIRT2 Ab and anti-His Ab or anti-GST Ab, confirmed the identity of the protein. SIRT2 activity studies showed that the SIRT2 construct SIRT2.6 (residues 34-389) was the most active of the four constructs. This difference in activity could be due to the presence of fusion tags or the truncation of the protein. The C-terminal fragment of SIRT2 (residues 356 – 389) have not been crystallised to date and may be involved in the activity of the protein as observed in other sirtuins (Davenport *et al.*, 2014).

Substrate and inhibitor interactions with sirtuins have been known to result in dynamic structural changes. Recently, inhibitor binding was shown to induce structural changes of SIRT2 resulting in the formation of a closed conformation on comparison with the apo-SIRT2 structure (Yamagata *et al.*, 2014). Additionally, the co-factor binding loop was shown to change in close relation to the catalytic events which comprise the formation of a covalent 1'-O-alkylamidate intermediate between the two substrates (NAD^+ and acetyl lysine) under nicotinamide release (Sanders *et al.*, 2010). This loop is in close proximity to the proposed region for the C-terminal flanking extensions (residues 356-389) and may also be involved in the catalytic activity of SIRT2 [Figure 5.4.1]. As a result of this, removal of the C-terminal flanking region (residues 356-389) would result in a decrease in SIRT2 activity as observed in constructs SIRT2.1,

SIRT2.4 and SIRT2.7 (residues 34-356). The crystal structure of SIRT2.6 or other full length constructs would help in understanding the role of the C-terminal flanking region in the catalytic activity of SIRT2.

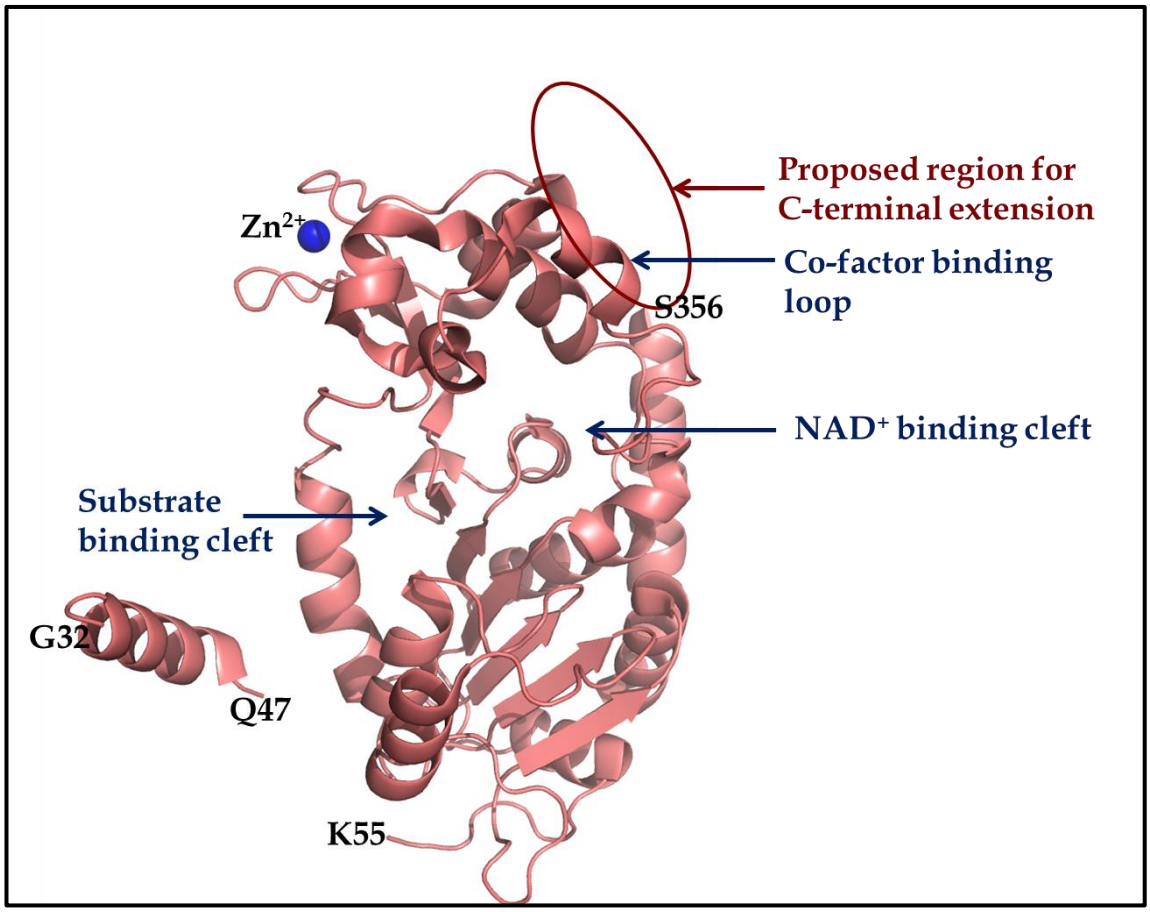


Figure 5.4. 1: Crystal structure of apo-form of SIRT2.

(PDB code: 3ZGO) The proposed region for the SIRT2 C-terminal flanking extension is indicated in the figure.

SIRT7

Three constructs of full length SIRT7 with a His-tag or a GST-tag were expressed and purified for crystallisation trials [Table 5.2.2]. Soluble expression of the protein was initially achieved using a combination of larger affinity tags, such as the GST-tag, and *E. coli* ArcticExpress cells which aid slow expression of the protein at lower temperatures to ensure correct refolding and enhance solubility of the protein. However, precipitation of SIRT7.6 due to cleavage of the tag prompted the design of construct SIRT7.7, which took advantage of a different vector and protease for tag cleavage. Untagged SIRT7 was successfully

obtained on cleavage of the GST-tag from purified SIRT7.7. The differences in the linker region between pGEx-6P-1 and pET-41 Ek/LIC may also contribute to the cleaved SIRT7.7 remaining soluble following cleavage. Mass spectrometric analysis and western blot analysis using anti-His or anti-GST Ab confirmed the identity of the protein. As insoluble expression of SIRT7.1a was observed, its retention of NAD^+ binding activity following refolding was confirmed using the SIRT7 activity assay. This assay was performed with biotinylated-NAD^+ and the results analysed via a modified Western blot.

SIRT1

Various constructs of SIRT1 were designed for expression and purification of SIRT1. However, difficulties in expression of this protein were encountered. The difference in primary and secondary structure, such as length of loops and longer terminal flanking regions, of SIRT1 compared to other members of the sirtuin family may be a contributing factor for these results. Further trials such as eukaryotic expression of this protein or its expression using other affinity tags such as SUMO or MBP-tags may yield favourable results.

5.4.2. Protein Crystallisation

Protein crystallographers are often confronted with recalcitrant proteins that are not readily crystallisable, or which crystallise in problematic forms. A general problem in protein crystallisation is that the crystallisation conditions required for crystallisation of a given protein, which include a combination of pH, ionic strength, temperature, protein concentration, the presence of various salts, ligands or additives, the type of precipitant and the actual crystallization method to use (hanging drop, sitting drop, dialysis, etc.), are practically impossible to predict in advance. Additionally, minor changes in these factors could result in non-reproducible the crystallisation results.

Nucleation and crystal growth are affected by chemical and physical parameters, as well as the crystallisation method used. A wide range of different conditions based on various screening kits were used in attempts to

crystallise SIRT2 and SIRT7. Almost all conditions tried, resulted in clear drops, amorphous precipitate or phase separation. A few conditions yielded crystalline precipitate and potential crystallisation hits for both SIRT2 and SIRT7 were obtained. In spite of varying several parameters in these conditions, the attempts to crystallise both proteins to produce diffraction quality crystals were unsuccessful.

Despite the inability to arrive at the right conditions, some basic observations were made based on the trials so far. The protein seems to prefer lower temperatures rather than room temperature or higher for incubation. At temperatures above 10 ^{o}C, lower concentrations (below 5 mg/mL) of protein ensured it remained in solution and heavy precipitation was minimised. It was also observed that lower pH at any concentration caused turbidity of the protein solution in the drop over a period of time.

One of the ways to overcome the obstacle of crystal growth is to set up crystallisation trials of the proteins with their affinity tags as this may contribute to crystal contacts and nucleation of the protein. Additionally, conformational heterogeneity is introduced by the flexible linker between the affinity tag and the protein of interest. To this end, crystallisation trials for SIRT2.7, SIRT7.6a and SIRT7.7 were set up with their respective affinity tags. Activity assays for SIRT2 and SIRT7 were also performed to confirm that the final purified protein used for crystallisation was active. However, no diffraction quality crystals were obtained. This could be due to the intrinsic disorder of sirtuins which may be overcome by co-crystallisation of the proteins with their respective substrates or ligands.

The inability of SIRT7.6a crystals to diffract may be indicative of its ability to nucleate and crystallise. However, optimisation of these crystallisation conditions would be required to yield diffraction quality crystals. If future trials also result in crystals that do not diffract, optimisation experiments to help in the formation of ordered crystals would be required. This is because the internal order of the crystals is crucial in protein crystallography and not the

outward appearance alone. Often the two go hand in hand but a true assessment of a crystal's quality can only be determined by X-ray diffraction. The minor variations in protein preparation and crystallisation may contribute to non-reproducibility of published SIRT2 crystals as observed in this study. These differences may include variation in the chemicals used due to purchase from different manufacturer's or simply the environment in which the experiments were performed. Idiosyncratic variation of individual procedures would also contribute to the inability of the protein to crystallise.

Currently, there are numerous 'uncrystallisable' proteins but, as molecular biology methodologies improve and the knowledge of crystallisation advances, these unyielding cases may be solved. With respect to sirtuins, there are still many approaches that can be explored to attain diffraction quality crystals.

These approaches include changes in both protein preparation and protein crystallisation techniques. One of the routes for protein preparation includes variation of mammalian sirtuin expression hosts such as mammalian, or insect cells through incorporation of the gene of interest in a baculovirus vector. These methods ensure a high level of protein expression with the refolding and post-translational modifications, an ability that bacterial hosts lack. However, these methods were beyond the scope of this study. Another approach would be to modify the expression tags to enable the expression of mammalian sirtuins covalently linked to other tags such as cleavable SUMO-His, thioredoxin or MBP-tags. These tags aid in soluble expression of the protein and can help in affinity purification of the protein. Modification to bio-macromolecular crystal growth techniques by using methods such as sandwich drop, free interface diffusion or microdialysis may help crystallisation of sirtuins.

Over the past few decades, much work in the field of sirtuins has been devoted to SIRT1 and SIRT2. However, the interplay of sirtuins demands that we deepen our understanding of each sirtuin to help modulate their activity such that the effect is beneficial to our health rather than it being detrimental (Cen *et al.*, 2011).

The known crystal structures of sirtuins have helped us build limited knowledge of the structural requirements for sirtuin activation and inhibition, however elucidating the full length structures of all sirtuins would greatly potentiate this study and help its effective modulation.

Such studies, can be used to identify regions of the protein responsible for its function by comparing it with proteins that are structurally similar and yet functionally different (as seen among members of the sirtuin family). Elucidating the full length structures of SIRT1, SIRT2 and SIRT7, will serve as a starting point for a rational program of experimentation such as site directed mutagenesis, ligand binding studies, rational drug design and molecular dynamic studies. This might lead to additional therapeutic opportunities as well as refinement of current strategies that are aimed at ageing and ageing related diseases.

REFERENCES

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., Mccoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-221.

Afshar, G. and Murnane, J. P. 1999. Characterization of a human gene with sequence homology to *Saccharomyces cerevisiae* SIR2. *Gene,* 234**,** 161-168.

Ahmad, N., Gabius, H. J., Andre, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F. and Brewer, C. F. 2004. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *J Biol Chem,* 279**,** 10841-10847.

Alcain, F. J. and Villalba, J. M. 2009a. Sirtuin activators. *Expert Opin Ther Pat,* 19**,** 403-414.

Alcain, F. J. and Villalba, J. M. 2009b. Sirtuin inhibitors. *Expert Opin Ther Pat,* 19**,** 283-294.

Alexander, C. and Rietschel, E. T. 2001. Bacterial lipopolysaccharides and innate immunity. *J Endotoxin Res,* 7**,** 167-202.

Allison, S. J. and Milner, J. 2007. SIRT3 is pro-apoptotic and participates in distinct basal apoptotic pathways. *Cell Cycle,* 6**,** 2669-2677.

Anderson, R. M., Bitterman, K. J., Wood, J. G., Medvedik, O. and Sinclair, D. A. 2003. Nicotinamide and PNC1 govern lifespan extension by calorie restriction in *Saccharomyces cerevisiae*. *Nature,* 423**,** 181-185.

Andre, S., Ortega, P. J., Perez, M. A., Roy, R. and Gabius, H. J. 1999. Lactose-containing starburst dendrimers: influence of dendrimer generation and binding-site orientation of receptors (plant/animal lectins and immunoglobulins) on binding properties. *Glycobiology,* 9**,** 1253-1261.

Andre, S., Pieters, R. J., Vrasidas, I., Kaltner, H., Kuwabara, I., Liu, F. T., Liskamp, R. M. and Gabius, H. J. 2001. Wedgelike glycodendrimers as inhibitors of binding of mammalian galectins to glycoproteins, lactose maxiclusters, and cell surface glycoconjugates. *Chembiochem,* 2**,** 822-830.

Asher, G., Gatfield, D., Stratmann, M., Reinke, H., Dibner, C., Kreppel, F., Mostoslavsky, R., Alt, F. W. and Schibler, U. 2008. SIRT1 regulates circadian clock gene expression through PER2 deacetylation. *Cell,* 134**,** 317-328.

Ashraf, N., Zino, S., Macintyre, A., Kingsmore, D., Payne, A. P., George, W. D. and Shiels, P. G. 2006. Altered sirtuin expression is associated with node-positive breast cancer. *Br J Cancer,* 95**,** 1056-1061.

Aub, J. C., Sanford, B. H. and Cote, M. N. 1965. Studies on reactivity of tumor and normal cells to a wheat germ agglutinin. *Proc Natl Acad Sci U S A,* 54**,** 396-399.

Aub, J. C., Tieslau, C. and Lankester, A. 1963. Reactions of Normal and Tumor Cell Surfaces to Enzymes. I. Wheat-Germ Lipase and Associated Mucopolysaccharides. *Proc Natl Acad Sci U S A,* 50**,** 613-619.

Bae, N. S., Swanson, M. J., Vassilev, A. and Howard, B. H. 2004. Human histone deacetylase SIRT2 interacts with the homeobox transcription factor HOXA10. *J Biochem,* 135**,** 695-700.

Banerjee, K. K., Ayyub, C., Ali, S. Z., Mandot, V., Prasad, N. G. and Kolthur-Seetharam, U. 2012. dSir2 in the adult fat body, but not in muscles, regulates life span in a diet-dependent manner. *Cell Rep,* 2**,** 1485-1491.

Baneyx, F. 1999. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol,* 10**,** 411-421.

Barber, M. F., Michishita-Kioi, E., Xi, Y., Tasselli, L., Kioi, M., Moqtaderi, Z., Tennen, R. I., Paredes, S., Young, N. L., Chen, K., Struhl, K., Garcia, B. A., Gozani, O., Li, W. and Chua, K. F. 2012. SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation. *Nature,* 487**,** 114-118.

Barondes, S. H., Castronovo, V., Cooper, D. N., Cummings, R. D., Drickamer, K., Feizi, T., Gitt, M. A., Hirabayashi, J., Hughes, C., Kasai, K. and Et Al. 1994a. Galectins: a family of animal β-galactoside-binding lectins. *Cell,* 76**,** 597-598.

Barondes, S. H., Cooper, D. N., Gitt, M. A. and Leffler, H. 1994b. Galectins. Structure and function of a large family of animal lectins. *J Biol Chem,* 269**,** 20807-20810.

Bartova, E., Krejci, J., Harnicarova, A., Galiova, G. and Kozubek, S. 2008. Histone modifications and nuclear architecture: a review. *J Histochem Cytochem,* 56**,** 711-721.

Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr,* 67**,** 271-281.

Baur, J. A., Pearson, K. J., Price, N. L., Jamieson, H. A., Lerin, C., Kalra, A., Prabhu, V. V., Allard, J. S., Lopez-Lluch, G., Lewis, K., Pistell, P. J., Poosala, S., Becker, K. G., Boss, O., Gwinn, D., Wang, M., Ramaswamy, S., Fishbein, K. W., Spencer, R. G., Lakatta, E. G., Le Couteur, D., Shaw, R. J., Navas, P., Puigserver, P., Ingram, D. K., De Cabo, R. and Sinclair, D. A. 2006. Resveratrol improves health and survival of mice on a high-calorie diet. *Nature,* 444**,** 337-342.

Bawono, P. and Heringa, J. 2014. PRALINE: a versatile multiple sequence alignment toolkit. *Methods Mol Biol,* 1079**,** 245-262.

Bedalov, A., Gatbonton, T., Irvine, W. P., Gottschling, D. E. and Simon, J. A. 2001. Identification of a small molecule inhibitor of Sir2p. *Proc Natl Acad Sci U S A,* 98**,** 15113-15118.

Bemmo, A., Dias, C., Rose, A. A., Russo, C., Siegel, P. and Majewski, J. 2010. Exon-level transcriptome profiling in murine breast cancer reveals splicing changes specific to tumors with different metastatic abilities. *PLoS One,* 5**,** e11981.

Bernerd, F., Sarasin, A. and Magnaldo, T. 1999. Galectin-7 overexpression is associated with the apoptotic process in UVB-induced sunburn keratinocytes. *Proc Natl Acad Sci U S A,* 96**,** 11329-11334.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol,* 112**,** 535-542.

Bharathi, S. S., Zhang, Y., Mohsen, A. W., Uppala, R., Balasubramani, M., Schreiber, E., Uechi, G., Beck, M. E., Rardin, M. J., Vockley, J., Verdin, E., Gibson, B. W., Hirschey, M. D. and Goetzman, E. S. 2013. Sirtuin 3 (SIRT3) protein regulates long-chain acyl-CoA dehydrogenase by deacetylating conserved lysines near the active site. *J Biol Chem,* 288**,** 33837-33847.

Blander, G. and Guarente, L. 2004. The Sir2 family of protein deacetylases. *Annu Rev Biochem,* 73**,** 417-435.

Blundell, T. L. and Johnson, L. N. 1976. *Protein crystallography*, Academic Press, London, New York, San Francisco.

Boas, U., Christensen, J. B. and Heegaard, P. M. H. 2006. *Dendrimers in Medicine and Biotechnology: New Molecular Tools,* Cambridge, United Kingdom, RSC Publishing.

Bosch-Presegue, L. and Vaquero, A. 2011. The dual role of sirtuins in cancer. *Genes Cancer,* 2**,** 648-662.

Bourne, Y., Bolgiano, B., Liao, D. I., Strecker, G., Cantau, P., Herzberg, O., Feizi, T. and Cambillau, C. 1994. Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. *Nat Struct Biol,* 1**,** 863-870.

Boyd, W. C. and Shapleigh, E. 1954. Specific Precipitating Activity of Plant Agglutinins (Lectins). *Science,* 119**,** 419.

Brachmann, C. B., Sherman, J. M., Devine, S. E., Cameron, E. E., Pillus, L. and Boeke, J. D. 1995. The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. *Genes Dev,* 9**,** 2888-2902.

Bradbury, C. A., Khanim, F. L., Hayden, R., Bunce, C. M., White, D. A., Drayson, M. T., Craddock, C. and Turner, B. M. 2005. Histone deacetylases in acute myeloid leukaemia show a distinctive pattern of expression that changes selectively in response to deacetylase inhibitors. *Leukemia,* 19**,** 1751-1759.

Bragg, W. L. 1913. The Structure of Some Crystals as Indicated by Their Diffraction of X-rays. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character,* 89**,** 248-277.

Branderhorst, H. M., Liskamp, R. M., Visser, G. M. and Pieters, R. J. 2007. Strong inhibition of cholera toxin binding by galactose dendrimers. *Chem Commun (Camb)***,** 5043-5045.

Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D. and Broach, J. R. 1993. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev,* 7**,** 592-604.

Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M. and Broach, J. R. 1996. Efficient transcriptional silencing in *Saccharomyces cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol Cell Biol,* 16**,** 4349-4356.

Brewer, C. F. 2002. Binding and cross-linking properties of galectins. *Biochim Biophys Acta,* 1572**,** 255-262.

Brewer, C. F. 2004. Thermodynamic binding studies of galectin-1, -3 and -7. *Glycoconj J,* 19**,** 459-465.

Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y., Tran, H., Ross, S. E., Mostoslavsky, R., Cohen, H. Y., Hu, L. S., Cheng, H. L., Jedrychowski, M. P., Gygi, S. P., Sinclair, D. A., Alt, F. W. and Greenberg, M.

E. 2004. Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science,* 303**,** 2011-2015.

Brunger, A. T. 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* 355**,** 472-475.

Bryk, M., Banerjee, M., Murphy, M., Knudsen, K. E., Garfinkel, D. J. and Curcio, M. J. 1997. Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev,* 11**,** 255-269.

Bucher, M. H., Evdokimov, A. G. and Waugh, D. S. 2002. Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein. *Acta Crystallogr D Biol Crystallogr,* 58**,** 392-397.

Burnett, C., Valentini, S., Cabreiro, F., Goss, M., Somogyvari, M., Piper, M. D., Hoddinott, M., Sutphin, G. L., Leko, V., Mcelwee, J. J., Vazquez-Manrique, R. P., Orfila, A. M., Ackerman, D., Au, C., Vinti, G., Riesen, M., Howard, K., Neri, C., Bedalov, A., Kaeberlein, M., Soti, C., Partridge, L. and Gems, D. 2011. Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and Drosophila. *Nature,* 477**,** 482-485.

Califice, S., Castronovo, V. and Van Den Brule, F. 2004. Galectin-3 and cancer. *Int J Oncol,* 25**,** 983-992.

Camby, I., Le Mercier, M., Lefranc, F. and Kiss, R. 2006. Galectin-1: a small protein with major functions. *Glycobiology,* 16**,** 137r-157r.

Campion, C. G., Labrie, M., Lavoie, G. and St-Pierre, Y. 2013. Expression of galectin-7 is induced in breast cancer cells by mutant p53. *PLoS One,* 8**,** e72468.

Canto, C. and Auwerx, J. 2012. Targeting sirtuin 1 to improve metabolism: all you need is NAD^+? *Pharmacol Rev,* 64**,** 166-187.

Cao, Z., Said, N., Amin, S., Wu, H. K., Bruce, A., Garate, M., Hsu, D. K., Kuwabara, I., Liu, F. T. and Panjwani, N. 2002a. Galectins-3 and -7, but not galectin-1, play a role in re-epithelialization of wounds. *J Biol Chem,* 277**,** 42299-42305.

Cao, Z., Said, N., Wu, H. K., Kuwabara, I., Liu, F. T. and Panjwani, N. 2003. Galectin-7 as a potential mediator of corneal epithelial cell migration. *Arch Ophthalmol,* 121**,** 82-86.

Cao, Z., Wu, H. K., Bruce, A., Wollenberg, K. and Panjwani, N. 2002b. Detection of differentially expressed genes in healing mouse corneas, using cDNA microarrays. *Invest Ophthalmol Vis Sci,* 43**,** 2897-2904.

Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T. and Ben-Tal, N. 2013. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Isr J Chem,* 53**,** 199-206.

Cen, Y., Falco, J. N., Xu, P., Youn, D. Y. and Sauve, A. A. 2011. Mechanism-based affinity capture of sirtuins. *Org Biomol Chem,* 9**,** 987-993.

Cerri, D. G., Rodrigues, L. C., Stowell, S. R., Araujo, D. D., Coelho, M. C., Oliveira, S. R., Bizario, J. C., Cummings, R. D., Dias-Baruffi, M. and Costa, M. C. 2008. Degeneration of dystrophic or injured skeletal muscles induces high expression of Galectin-1. *Glycobiology,* 18**,** 842-850.

Chabre, Y. M. and Roy, R. 2012. Dendrimer-Coated Carbohydrate Residues as Drug Delivery Trojan Horses in Glycoscience. *Dendrimer-Based Drug Delivery Systems.* John Wiley & Sons, Inc.

Chakrabarti, P., English, T., Karki, S., Qiang, L., Tao, R., Kim, J., Luo, Z., Farmer, S. R. and Kandror, K. V. 2011. SIRT1 controls lipolysis in adipocytes via FOXO1-mediated expression of ATGL. *J Lipid Res,* 52**,** 1693-1701.

Chang-Hong, R., Wada, M., Koyama, S., Kimura, H., Arawaka, S., Kawanami, T., Kurita, K., Kadoya, T., Aoki, M., Itoyama, Y. and Kato, T. 2005. Neuroprotective effect of oxidized galectin-1 in a transgenic mouse model of amyotrophic lateral sclerosis. *Exp Neurol,* 194**,** 203-211.

Chang, J. H., Kim, H. C., Hwang, K. Y., Lee, J. W., Jackson, S. P., Bell, S. D. and Cho, Y. 2002. Structural basis for the NAD-dependent deacetylase mechanism of Sir2. *J Biol Chem,* 277**,** 34489-34498.

Chen, S., Seiler, J., Santiago-Reichelt, M., Felbel, K., Grummt, I. and Voit, R. 2013. Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Mol Cell,* 52**,** 303-313.

Chen, V. B., Arendall, W. B. R., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr,* 66**,** 12-21.

Cooper, D. N. 2002. Galectinomics: finding themes in complexity. *Biochim Biophys Acta,* 1572**,** 209-231.

Couraud, P. O., Casentini-Borocz, D., Bringman, T. S., Griffith, J., Mcgrogan, M. and Nedwin, G. E. 1989. Molecular cloning, characterization, and expression of a human 14-kDa lectin. *J Biol Chem,* 264**,** 1310-1316.

Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., Van Den Berg, T. K., Van Der Merwe, P. A., Watt, S. M. and Varki, A. 1998. Siglecs: a family of sialic-acid binding lectins. *Glycobiology,* 8**,** v.

Cummings, R. D. and Etzler, M. E. 2009. *R-type Lectins,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Cummings, R. D. and Liu, F. T. 2009. *Galectins*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Cummings, R. D. and Mcever, R. P. 2009. *C-type Lectins,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

D'arcy, A., Villard, F. and Marsh, M. 2007. An automated microseed matrix-screening method for protein crystallization. *Acta Crystallogr D Biol Crystallogr,* 63**,** 550-554.

Dan, L., Klimenkova, O., Klimiankou, M., Klusman, J. H., Van Den Heuvel-Eibrink, M. M., Reinhardt, D., Welte, K. and Skokowa, J. 2012. The role of sirtuin 2 activation by nicotinamide phosphoribosyltransferase in the aberrant proliferation and survival of myeloid leukemia cells. *Haematologica,* 97**,** 551-559.

Danguy, A., Camby, I. and Kiss, R. 2002. Galectins and cancer. *Biochim Biophys Acta,* 1572**,** 285-293.

Davenport, A. M., Huber, F. M. and Hoelz, A. 2014. Structural and functional analysis of human SIRT1. *J Mol Biol,* 426**,** 526-541.

De Nigris, F., Cerutti, J., Morelli, C., Califano, D., Chiariotti, L., Viglietto, G., Santelli, G. and Fusco, A. 2002. Isolation of a SIR-like gene, SIR-T8, that is overexpressed in thyroid carcinoma cell lines and tissues. *Br J Cancer,* 86**,** 917-923.

De Oliveira, R. M., Sarkander, J., Kazantsev, A. G. and Outeiro, T. F. 2012. SIRT2 as a Therapeutic Target for Age-Related Disorders. *Front Pharmacol,* 3**,** 82.

De Waard, A., Hickman, S. and Kornfeld, S. 1976. Isolation and properties of β-galactoside binding lectins of calf heart and lung. *J Biol Chem,* 251**,** 7581-7587.

Delacour, D., Koch, A. and Jacob, R. 2009. The role of galectins in protein trafficking. *Traffic,* 10**,** 1405-1413.

Demers, M., Biron-Pain, K., Hebert, J., Lamarre, A., Magnaldo, T. and St-Pierre, Y. 2007. Galectin-7 in lymphoma: elevated expression in human lymphoid malignancies and decreased lymphoma dissemination by antisense strategies in experimental model. *Cancer Res,* 67**,** 2824-2829.

Demers, M., Magnaldo, T. and St-Pierre, Y. 2005. A novel function for galectin-7: promoting tumorigenesis by up-regulating MMP-9 gene expression. *Cancer Res,* 65**,** 5205-5210.

Demers, M., Rose, A. A., Grosset, A. A., Biron-Pain, K., Gaboury, L., Siegel, P. M. and St-Pierre, Y. 2010. Overexpression of galectin-7, a myoepithelial cell marker, enhances spontaneous metastasis of breast cancer cells. *Am J Pathol,* 176**,** 3023-3031.

Dimick, S. M., Powell, S. C., Mcmahon, S. A., Moothoo, D. N., Naismith, J. H. and Toone, E. J. 1999. On the Meaning of Affinity: Cluster Glycoside Effects and Concanavalin A. *J Am Chem Soc,* 121**,** 10286-10296.

Disch, J. S., Evindar, G., Chiu, C. H., Blum, C. A., Dai, H., Jin, L., Schuman, E., Lind, K. E., Belyanskaya, S. L., Deng, J., Coppo, F., Aquilani, L., Graybill, T. L., Cuozzo, J. W., Lavu, S., Mao, C., Vlasuk, G. P. and Perni, R. B. 2013. Discovery of thieno[3,2-d]pyrimidine-6-carboxamides as potent inhibitors of SIRT1, SIRT2, and SIRT3. *J Med Chem,* 56**,** 3666-3679.

Donmez, G., Arun, A., Chung, C. Y., Mclean, P. J., Lindquist, S. and Guarente, L. 2012. SIRT1 protects against α-synuclein aggregation by activating molecular chaperones. *J Neurosci,* 32**,** 124-132.

Donmez, G., Wang, D., Cohen, D. E. and Guarente, L. 2010. SIRT1 suppresses β-amyloid production by activating the α-secretase gene ADAM10. *Cell,* 142**,** 320-332.

Drickamer, K. 1988. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J Biol Chem,* 263**,** 9557-9560.

Drickamer, K. and Taylor, M. E. 1993. Biology of animal lectins. *Annu Rev Cell Biol,* 9**,** 237-264.

Drickamer, K. and Taylor, M. E. 1998. Evolving views of protein glycosylation. *Trends Biochem Sci,* 23**,** 321-324.

Dryden, S. C., Nahhas, F. A., Nowak, J. E., Goustin, A. S. and Tainsky, M. A. 2003. Role for human SIRT2 NAD-dependent deacetylase activity in control of mitotic exit in the cell cycle. *Mol Cell Biol,* 23**,** 3173-3185.

Du, J., Jiang, H. and Lin, H. 2009. Investigating the ADP-ribosyltransferase activity of sirtuins with NAD analogues and 32P-NAD. *Biochemistry,* 48**,** 2878-2890.

Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., Choi, B. H., He, B., Chen, W., Zhang, S., Cerione, R. A., Auwerx, J., Hao, Q. and Lin, H. 2011. Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science,* 334**,** 806-809.

Eisenberg, D., Schwarz, E., Komaromy, M. and Wall, R. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol,* 179**,** 125-142.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

Ermakova, E., Miller, M. C., Nesmelova, I. V., Lopez-Merino, L., Berbis, M. A., Nesmelov, Y., Tkachev, Y. V., Lagartera, L., Daragan, V. A., Andre, S., Canada, F. J., Jimenez-Barbero, J., Solis, D., Gabius, H. J. and Mayo, K., H. 2013. Lactose binding to human galectin-7 (p53-induced gene 1) induces long-range effects through the protein resulting in increased dimer stability and evidence for positive cooperativity. *Glycobiology,* 23**,** 508-523.

Espada, J., Ballestar, E., Santoro, R., Fraga, M. F., Villar-Garea, A., Nemeth, A., Lopez-Serra, L., Ropero, S., Aranda, A., Orozco, H., Moreno, V., Juarranz, A., Stockert, J. C., Langst, G., Grummt, I., Bickmore, W. and Esteller, M. 2007. Epigenetic disruption of ribosomal RNA genes and nucleolar architecture in DNA methyltransferase 1 (Dnmt1) deficient cells. *Nucleic Acids Res,* 35**,** 2191-2198.

Evans, J., Yap, J., Gamage, T., Salamonsen, L., Dimitriadis, E. and Menkhorst, E. 2014. Galectin-7 is important for normal uterine repair following menstruation. *Mol Hum Reprod*.

Evans, P. R. and Murshudov, G. N. 2013. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr,* 69**,** 1204-1214.

Feldman, J. L., Baeza, J. and Denu, J. M. 2013. Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins. *J Biol Chem,* 288**,** 31350-31356.

Feldman, J. L., Dittenhafer-Reed, K. E. and Denu, J. M. 2012. Sirtuin catalysis and regulation. *J Biol Chem,* 287**,** 42419-42427.

Finley, L. W., Haas, W., Desquiret-Dumas, V., Wallace, D. C., Procaccio, V., Gygi, S. P. and Haigis, M. C. 2011. Succinate dehydrogenase is a direct target of sirtuin 3 deacetylase activity. *PLoS One,* 6**,** e23295.

Finnin, M. S., Donigian, J. R. and Pavletich, N. P. 2001. Structure of the histone deacetylase SIRT2. *Nat Struct Biol,* 8**,** 621-625.

Flick, F. and Luscher, B. 2012. Regulation of sirtuin function by posttranslational modifications. *Front Pharmacol,* 3**,** 29.

Ford, E., Voit, R., Liszt, G., Magin, C., Grummt, I. and Guarente, L. 2006. Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. *Genes Dev,* 20**,** 1075-1080.

Ford, J., Jiang, M. and Milner, J. 2005. Cancer-specific functions of SIRT1 enable human epithelial cancer cell growth and survival. *Cancer Res,* 65**,** 10457-10463.

Franz, H. 1988. The ricin story. *Adv. Lectin Res.,* 1**,** 10-25.

Fréchet, J. M. J. and Tomalia, D. A. *Dendrimers and Other Dendritic Polymers*, Wiley-VCH, Weinheim, Germany.

Frye, R. A. 1999. Characterization of five human cDNAs with homology to the yeast SIR2 gene: Sir2-like proteins (sirtuins) metabolize NAD and may have protein ADP-ribosyltransferase activity. *Biochem Biophys Res Commun,* 260**,** 273-279.

Frye, R. A. 2000. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem Biophys Res Commun,* 273**,** 793-798.

Gabius, H. J. 1991. Detection and functions of mammalian lectins--with emphasis on membrane lectins. *Biochim Biophys Acta,* 1071**,** 1-18.

Gabius, H. J. 2000. Biological information transfer beyond the genetic code: the sugar code. *Naturwissenschaften,* 87**,** 108-121.

Gardikis, K., Micha-Screttas, M., Demetzos, C. and Steele, B. R. 2012. Dendrimers and the development of new complex nanomaterials for biomedical applications. *Curr Med Chem,* 19**,** 4913-4928.

Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C. and Escalante-Semerena, J. C. 2007. N-lysine propionylation controls the activity of propionyl-CoA synthetase. *J Biol Chem,* 282**,** 30239-30245.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. 2005. *Protein Identification and Analysis Tools on the ExPASy Server*, Humana Press.

Gertz, M., Fischer, F., Nguyen, G. T., Lakshminarasimhan, M., Schutkowski, M., Weyand, M. and Steegborn, C. 2013. Ex-527 inhibits Sirtuins by exploiting their unique NAD^+-dependent deacetylation mechanism. *Proc Natl Acad Sci U S A,* 110**,** E2772-2781.

Gertz, M., Nguyen, G. T., Fischer, F., Suenkel, B., Schlicker, C., Franzel, B., Tomaschewski, J., Aladini, F., Becker, C., Wolters, D. and Steegborn, C. 2012. A molecular mechanism for direct sirtuin activation by resveratrol. *PLoS One,* 7**,** e49761.

Goldring, K., Jones, G. E., Thiagarajah, R. and Watt, D. J. 2002. The effect of galectin-1 on the differentiation of fibroblasts and myoblasts *in vitro*. *J Cell Sci,* 115**,** 355-366.

Goodman, C. K., Wolfenden, M. L., Nangia-Makker, P., Michel, A. K., Raz, A. and Cloninger, M. J. 2014. Multivalent scaffolds induce galectin-3 aggregation into nanoparticles. *Beilstein J Org Chem,* 10**,** 1570-1577.

Gorityala, B. K., Lu, Z., Leow, M. L., Ma, J. and Liu, X. W. 2012. Design of a "turn-off/turn-on" biosensor: understanding carbohydrate-lectin interactions for use in noncovalent drug delivery. *J Am Chem Soc,* 134**,** 15229-15232.

Gottschling, D. E., Aparicio, O. M., Billington, B. L. and Zakian, V. A. 1990. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell,* 63**,** 751-762.

Green, D. W., Ingram, V. M. and Perutz, M. F. 1954. The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proc R Soc A Math Phys Sci,* 225**,** 287-307.

Guarente, L. 2011. Franklin H. Epstein Lecture: Sirtuins, aging, and medicine. *N Engl J Med,* 364**,** 2235-2244.

Guarente, L. 2013. Calorie restriction and sirtuins revisited. *Genes Dev,* 27**,** 2072-2085.

Haigis, M. C. and Guarente, L. P. 2006. Mammalian sirtuins--emerging roles in physiology, aging, and calorie restriction. *Genes Dev,* 20**,** 2913-2921.

Haigis, M. C. and Sinclair, D. A. 2010. Mammalian sirtuins: biological insights and disease relevance. *Annu Rev Pathol,* 5**,** 253-295.

Hallows, W. C., Yu, W., Smith, B. C., Devries, M. K., Ellinger, J. J., Someya, S., Shortreed, M. R., Prolla, T., Markley, J. L., Smith, L. M., Zhao, S., Guan, K. L. and Denu, J. M. 2011. Sirt3 promotes the urea cycle and fatty acid oxidation during dietary restriction. *Mol Cell,* 41**,** 139-149.

Han, Y., Jin, Y. H., Kim, Y. J., Kang, B. Y., Choi, H. J., Kim, D. W., Yeo, C. Y. and Lee, K. Y. 2008. Acetylation of Sirt2 by p300 attenuates its deacetylase activity. *Biochem Biophys Res Commun,* 375**,** 576-580.

Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M. and Grunstein, M. 1995. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell,* 80**,** 583-592.

Hein, N., Hannan, K. M., George, A. J., Sanij, E. and Hannan, R. D. 2013. The nucleolus: an emerging target for cancer therapy. *Trends Mol Med,* 19**,** 643-654.

Heltweg, B., Gatbonton, T., Schuler, A. D., Posakony, J., Li, H., Goehle, S., Kollipara, R., Depinho, R. A., Gu, Y., Simon, J. A. and Bedalov, A. 2006. Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes. *Cancer Res,* 66**,** 4368-4377.

Hendrickson, W. A., Smith, J. L. and Sheriff, S. 1985. Direct phase determination based on anomalous scattering. *Methods Enzymol,* 115**,** 41-55.

Herranz, D., Munoz-Martin, M., Canamero, M., Mulero, F., Martinez-Pastor, B., Fernandez-Capetillo, O. and Serrano, M. 2010. Sirt1 improves healthy ageing and protects from metabolic syndrome-associated cancer. *Nat Commun,* 1**,** 1-8.

Herskovits, A. Z. and Guarente, L. 2014. SIRT1 in neurodevelopment and brain senescence. *Neuron,* 81**,** 471-483.

Hida, Y., Kubo, Y., Murao, K. and Arase, S. 2007. Strong expression of a longevity-related protein, SIRT1, in Bowen's disease. *Arch Dermatol Res,* 299**,** 103-106.

Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W. E., Yagi, F. and Kasai, K. 2002. Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim Biophys Acta,* 1572**,** 232-254.

Hirabayashi, J. and Kasai, K. 1993. The family of metazoan metal-independent β-galactoside-binding lectins: structure, function and molecular evolution. *Glycobiology,* 3**,** 297-304.

Hiratsuka, M., Inoue, T., Toda, T., Kimura, N., Shirayoshi, Y., Kamitani, H., Watanabe, T., Ohama, E., Tahimic, C. G., Kurimasa, A. and Oshimura, M. 2003. Proteomics-based identification of differentially expressed genes in human gliomas: down-regulation of SIRT2 gene. *Biochem Biophys Res Commun,* 309**,** 558-566.

Hirschey, M. D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D. B., Grueter, C. A., Harris, C., Biddinger, S., Ilkayeva, O. R., Stevens, R. D., Li, Y., Saha, A. K., Ruderman, N. B., Bain, J. R., Newgard, C. B., Farese, R. V., Jr., Alt, F. W., Kahn, C. R. and Verdin, E. 2010. SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature,* 464**,** 121-125.

Hoff, K. G., Avalos, J. L., Sens, K. and Wolberger, C. 2006. Insights into the sirtuin mechanism from ternary complexes containing NAD^+ and acetylated peptide. *Structure,* 14**,** 1231-1240.

Houtkooper, R. H., Pirinen, E. and Auwerx, J. 2012. Sirtuins as regulators of metabolism and healthspan. *Nat Rev Mol Cell Biol,* 13**,** 225-238.

Howitz, K. T., Bitterman, K. J., Cohen, H. Y., Lamming, D. W., Lavu, S., Wood, J. G., Zipkin, R. E., Chung, P., Kisielewski, A., Zhang, L. L., Scherer, B. and Sinclair, D. A. 2003. Small molecule activators of sirtuins extend Saccharomyces cerevisiae lifespan. *Nature,* 425**,** 191-196.

Hudgin, R. L., Pricer, W. E., Jr., Ashwell, G., Stockert, R. J. and Morell, A. G. 1974. The isolation and properties of a rabbit liver binding protein specific for asialoglycoproteins. *J Biol Chem,* 249**,** 5536-5543.

Huffman, D. M., Grizzle, W. E., Bamman, M. M., Kim, J. S., Eltoum, I. A., Elgavish, A. and Nagy, T. R. 2007. SIRT1 is significantly elevated in mouse and human prostate cancer. *Cancer Res,* 67**,** 6612-6618.

Hughes, R. C. 1999. Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim Biophys Acta,* 1473**,** 172-185.

Imai, S., Armstrong, C. M., Kaeberlein, M. and Guarente, L. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature,* 403**,** 795-800.

Imai, S. I. and Guarente, L. 2014. NAD and sirtuins in aging and disease. *Trends Cell Biol,* 24**,** 464-471.

Inagaki, Y., Higashi, K., Kushida, M., Hong, Y. Y., Nakao, S., Higashiyama, R., Moro, T., Itoh, J., Mikami, T., Kimura, T., Shiota, G., Kuwabara, I. and Okazaki, I. 2008. Hepatocyte growth factor suppresses profibrogenic signal transduction via nuclear export of Smad3 with galectin-7. *Gastroenterology,* 134**,** 1180-1190.

Ingrassia, L., Nshimyumukiza, P., Dewelle, J., Lefranc, F., Wlodarczak, L., Thomas, S., Dielie, G., Chiron, C., Zedde, C., Tisnes, P., Van Soest, R., Braekman, J. C., Darro, F. and Kiss, R. 2006. A lactosylated steroid contributes *in vivo* therapeutic benefits in experimental models of mouse lymphoma and human glioblastoma. *J Med Chem,* 49**,** 1800-1807.

Inoue, T., Hiratsuka, M., Osaki, M. and Oshimura, M. 2007. The molecular biology of mammalian SIRT proteins: SIRT2 in cell cycle regulation. *Cell Cycle,* 6**,** 1011-1018.

Ion, G., Fajka-Boja, R., Toth, G. K., Caron, M. and Monostori, E. 2005. Role of p56lck and ZAP70-mediated tyrosine phosphorylation in galectin-1-induced cell death. *Cell Death Differ,* 12**,** 1145-1147.

Jackson, M. D. and Denu, J. M. 2002. Structural identification of 2'- and 3'-O-acetyl-ADP-ribose as novel metabolites derived from the Sir2 family of β-NAD^+-dependent histone/protein deacetylases. *J Biol Chem,* 277**,** 18535-18544.

Jeong, H., Cohen, D. E., Cui, L., Supinski, A., Savas, J. N., Mazzulli, J. R., Yates, J. R., 3rd, Bordone, L., Guarente, L. and Krainc, D. 2012. Sirt1 mediates neuroprotection from mutant huntingtin by activation of the TORC1 and CREB transcriptional pathway. *Nat Med,* 18**,** 159-165.

Jeong, S. M., Xiao, C., Finley, L. W., Lahusen, T., Souza, A. L., Pierce, K., Li, Y. H., Wang, X., Laurent, G., German, N. J., Xu, X., Li, C., Wang, R. H., Lee, J.,

Csibi, A., Cerione, R., Blenis, J., Clish, C. B., Kimmelman, A., Deng, C. X. and Haigis, M. C. 2013. SIRT4 has tumor-suppressive activity and regulates the cellular metabolic response to DNA damage by inhibiting mitochondrial glutamine metabolism. *Cancer Cell,* 23**,** 450-463.

Jiang, M., Wang, J., Fu, J., Du, L., Jeong, H., West, T., Xiang, L., Peng, Q., Hou, Z., Cai, H., Seredenina, T., Arbez, N., Zhu, S., Sommers, K., Qian, J., Zhang, J., Mori, S., Yang, X. W., Tamashiro, K. L., Aja, S., Moran, T. H., Luthi-Carter, R., Martin, B., Maudsley, S., Mattson, M. P., Cichewicz, R. H., Ross, C. A., Holtzman, D. M., Krainc, D. and Duan, W. 2012. Neuroprotective role of Sirt1 in mammalian models of Huntington's disease through activation of multiple Sirt1 targets. *Nat Med,* 18**,** 153-158.

Jiang, W., Wang, S., Xiao, M., Lin, Y., Zhou, L., Lei, Q., Xiong, Y., Guan, K. L. and Zhao, S. 2011. Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. *Mol Cell,* 43**,** 33-44.

Jin, L., Wei, W., Jiang, Y., Peng, H., Cai, J., Mao, C., Dai, H., Choy, W., Bemis, J. E., Jirousek, M. R., Milne, J. C., Westphal, C. H. and Perni, R. B. 2009. Crystal structures of human SIRT3 displaying substrate-induced conformational changes. *J Biol Chem,* 284**,** 24394-24405.

Kabsch, W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr,* 66**,** 125-132.

Kadoya, T., Oyanagi, K., Kawakami, E., Hasegawa, M., Inagaki, Y., Sohma, Y. and Horie, H. 2005. Oxidized galectin-1 advances the functional recovery after peripheral nerve injury. *Neurosci Lett,* 380**,** 284-288.

Kaeberlein, M., Mcvey, M. and Guarente, L. 1999. The SIR2/3/4 complex and SIR2 alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes Dev,* 13**,** 2570-2580.

Kanfi, Y., Naiman, S., Amir, G., Peshti, V., Zinman, G., Nahum, L., Bar-Joseph, Z. and Cohen, H. Y. 2012. The sirtuin SIRT6 regulates lifespan in male mice. *Nature,* 483**,** 218-221.

Kanfi, Y., Peshti, V., Gil, R., Naiman, S., Nahum, L., Levin, E., Kronfeld-Schor, N. and Cohen, H. Y. 2010. SIRT6 protects against pathological damage caused by diet-induced obesity. *Aging Cell,* 9**,** 162-173.

Kang, H., Suh, J. Y., Jung, Y. S., Jung, J. W., Kim, M. K. and Chung, J. H. 2011. Peptide switch is essential for Sirt1 deacetylase activity. *Mol Cell,* 44**,** 203-213.

Karin, M. 2006. NF-κB and cancer: mechanisms and targets. *Mol Carcinog,* 45**,** 355-361.

Karle, J. and Hauptman, H. 1956. A theory of phase determination for the four types of non-centrosymmetric space groups 1P222, 2P22, 3P12, 3P22. *Acta Crystallogr,* 9**,** 635-651.

Kikkeri, R., Garcia-Rubio, I. and Seeberger, P. H. 2009. Ru(II)-carbohydrate dendrimers as photoinduced electron transfer lectin biosensors. *Chem Commun (Camb)***,** 235-237.

Kikkeri, R., Grunstein, D. and Seeberger, P. H. 2010a. Lectin biosensing using digital analysis of Ru(II)-glycodendrimers. *J Am Chem Soc,* 132**,** 10230-10232.

Kikkeri, R., Liu, X., Adibekian, A., Tsai, Y. H. and Seeberger, P. H. 2010b. Facile synthesis of size dependent Ru(ii)-carbohydrate dendrimers via click chemistry. *Chem Commun (Camb),* 46**,** 2197-2199.

Kim, E. J., Kho, J. H., Kang, M. R. and Um, S. J. 2007. Active regulator of SIRT1 cooperates with SIRT1 and facilitates suppression of p53 activity. *Mol Cell,* 28**,** 277-290.

Kim, H. J., Jeon, H. K., Lee, J. K., Sung, C. O., Do, I. G., Choi, C. H., Kim, T. J., Kim, B. G., Bae, D. S. and Lee, J. W. 2013a. Clinical significance of galectin-7 in epithelial ovarian cancer. *Anticancer Res,* 33**,** 1555-1561.

Kim, H. S., Patel, K., Muldoon-Jacobs, K., Bisht, K. S., Aykin-Burns, N., Pennington, J. D., Van Der Meer, R., Nguyen, P., Savage, J., Owens, K. M., Vassilopoulos, A., Ozden, O., Park, S. H., Singh, K. K., Abdulkadir, S. A., Spitz, D. R., Deng, C. X. and Gius, D. 2010a. SIRT3 is a mitochondria-localized tumor suppressor required for maintenance of mitochondrial integrity and metabolism during stress. *Cancer Cell,* 17**,** 41-52.

Kim, H. S., Xiao, C., Wang, R. H., Lahusen, T., Xu, X., Vassilopoulos, A., Vazquez-Ortiz, G., Jeong, W. I., Park, O., Ki, S. H., Gao, B. and Deng, C. X. 2010b. Hepatic-specific disruption of SIRT6 in mice results in fatty liver formation due to enhanced glycolysis and triglyceride synthesis. *Cell Metab,* 12**,** 224-236.

Kim, J. E., Chen, J. and Lou, Z. 2008. DBC1 is a negative regulator of SIRT1. *Nature,* 451**,** 583-586.

Kim, J. K., Noh, J. H., Jung, K. H., Eun, J. W., Bae, H. J., Kim, M. G., Chang, Y. G., Shen, Q., Park, W. S., Lee, J. Y., Borlak, J. and Nam, S. W. 2013b. Sirtuin7 oncogenic potential in human hepatocellular carcinoma and its regulation by the tumor suppressors MiR-125a-5p and MiR-125b. *Hepatology,* 57**,** 1055-1067.

Kim, S. J., Hwang, J. A., Ro, J. Y., Lee, Y. S. and Chun, K. H. 2013c. Galectin-7 is epigenetically-regulated tumor suppressor in gastric cancer. *Oncotarget,* 4**,** 1461-1471.

Kojima, K., Ohhashi, R., Fujita, Y., Hamada, N., Akao, Y., Nozawa, Y., Deguchi, T. and Ito, M. 2008. A role for SIRT1 in cell growth and chemoresistance in prostate cancer PC3 and DU145 cells. *Biochem Biophys Res Commun,* 373**,** 423-428.

Kopitz, J., Andre, S., Von Reitzenstein, C., Versluis, K., Kaltner, H., Pieters, R. J., Wasano, K., Kuwabara, I., Liu, F. T., Cantz, M., Heck, A. J. and Gabius, H. J. 2003. Homodimeric galectin-7 (p53-induced gene 1) is a negative growth regulator for human neuroblastoma cells. *Oncogene,* 22**,** 6277-6288.

Krissinel, E. 2012. Enhanced fold recognition using efficient short fragment clustering. *J Mol Biol,* 1**,** 76-85.

Krissinel, E. and Henrick, K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2256-2268.

Krissinel, E. and Henrick, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol,* 372**,** 774-797.

Kuwabara, I., Kuwabara, Y., Yang, R. Y., Schuler, M., Green, D. R., Zuraw, B. L., Hsu, D. K. and Liu, F. T. 2002. Galectin-7 (PIG1) exhibits pro-apoptotic function through JNK activation and mitochondrial cytochrome c release. *J Biol Chem,* 277**,** 3487-3497.

Lain, S., Hollick, J. J., Campbell, J., Staples, O. D., Higgins, M., Aoubala, M., Mccarthy, A., Appleyard, V., Murray, K. E., Baker, L., Thompson, A., Mathers, J., Holland, S. J., Stark, M. J., Pass, G., Woods, J., Lane, D. P. and Westwood,

N. J. 2008. Discovery, *in vivo* activity, and mechanism of action of a small-molecule p53 activator. *Cancer Cell,* 13**,** 454-463.

Landry, J., Sutton, A., Tafrov, S. T., Heller, R. C., Stebbins, J., Pillus, L. and Sternglanz, R. 2000. The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc Natl Acad Sci U S A,* 97**,** 5807-5811.

Laskowski, R. A. and Swindells, M. B. 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model,* 51**,** 2778-2786.

Laurent, G., German, N. J., Saha, A. K., De Boer, V. C., Davies, M., Koves, T. R., Dephoure, N., Fischer, F., Boanca, G., Vaitheesvaran, B., Lovitch, S. B., Sharpe, A. H., Kurland, I. J., Steegborn, C., Gygi, S. P., Muoio, D. M., Ruderman, N. B. and Haigis, M. C. 2013. SIRT4 coordinates the balance between lipid synthesis and catabolism by repressing malonyl CoA decarboxylase. *Mol Cell,* 50**,** 686-698.

Leffler, H., Carlsson, S., Hedlund, M., Qian, Y. and Poirier, F. 2004. Introduction to galectins. *Glycoconj J,* 19**,** 433-440.

Lennerz, V., Fatho, M., Gentilini, C., Frye, R. A., Lifke, A., Ferel, D., Wolfel, C., Huber, C. and Wolfel, T. 2005. The response of autologous T cells to a human melanoma is dominated by mutated neoantigens. *Proc Natl Acad Sci U S A,* 102**,** 16013-16018.

Leonidas, D. D., Vatzaki, E. H., Vorum, H., Celis, J. E., Madsen, P. and Acharya, K. R. 1998. Structural basis for the recognition of carbohydrates by human galectin-7. *Biochemistry,* 37**,** 13930-13940.

Lepur, A., Salomonsson, E., Nilsson, U. J. and Leffler, H. 2012. Ligand induced galectin-3 protein self-association. *J Biol Chem,* 287**,** 21751-21756.

Leslie, A. G. 2006. The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr,* 62**,** 48-57.

Li, X., Zhang, S., Blander, G., Tse, J. G., Krieger, M. and Guarente, L. 2007. SIRT1 deacetylates and positively regulates the nuclear receptor LXR. *Mol Cell,* 28**,** 91-106.

Li, Y., Matsumori, H., Nakayama, Y., Osaki, M., Kojima, H., Kurimasa, A., Ito, H., Mori, S., Katoh, M., Oshimura, M. and Inoue, T. 2011. SIRT2 down-regulation in HeLa can induce p53 accumulation via p38 MAPK activation-dependent p300 decrease, eventually leading to apoptosis. *Genes Cells,* 16**,** 34-45.

Lin, S. J., Defossez, P. A. and Guarente, L. 2000. Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae. *Science,* 289**,** 2126-2128.

Lin, S. J., Kaeberlein, M., Andalis, A. A., Sturtz, L. A., Defossez, P. A., Culotta, V. C., Fink, G. R. and Guarente, L. 2002. Calorie restriction extends *Saccharomyces cerevisiae* lifespan by increasing respiration. *Nature,* 418**,** 344-348.

Lindstedt, R., Apodaca, G., Barondes, S. H., Mostov, K. E. and Leffler, H. 1993. Apical secretion of a cytosolic protein by Madin-Darby canine kidney cells. Evidence for polarized release of an endogenous lectin by a nonclassical secretory pathway. *J Biol Chem,* 268**,** 11750-11757.

Liszt, G., Ford, E., Kurtev, M. and Guarente, L. 2005. Mouse Sir2 homolog SIRT6 is a nuclear ADP-ribosyltransferase. *J Biol Chem,* 280**,** 21313-21320.

Liu, F. T., Patterson, R. J. and Wang, J. L. 2002. Intracellular functions of galectins. *Biochim Biophys Acta,* 1572**,** 263-273.

Liu, F. T. and Rabinovich, G. A. 2005. Galectins as modulators of tumour progression. *Nat Rev Cancer,* 5**,** 29-41.

Liu, X. Z. and Yan, D. 2007. Ageing and hearing loss. *J Pathol,* 211**,** 188-197.

Lok, D. J., Van Der Meer, P., De La Porte, P. W., Lipsic, E., Van Wijngaarden, J., Hillege, H. L. and Van Veldhuisen, D. J. 2010. Prognostic value of galectin-3, a novel marker of fibrosis, in patients with chronic heart failure: data from the DEAL-HF study. *Clin Res Cardiol,* 99**,** 323-328.

Lomb, D. J., Laurent, G. and Haigis, M. C. 2010. Sirtuins regulate key aspects of lipid metabolism. *Biochim Biophys Acta,* 1804**,** 1652-1657.

Lorber, B., Fischer, F., Bailly, M., Roy, H. and Kern, D. 2012. Protein analysis by dynamic light scattering: methods and techniques for students. *Biochem Mol Biol Educ,* 40**,** 372-382.

Loser, K., Sturm, A., Voskort, M., Kupas, V., Balkow, S., Auriemma, M., Sternemann, C., Dignass, A. U., Luger, T. A. and Beissert, S. 2009. Galectin-2 suppresses contact allergy by inducing apoptosis in activated CD8+ T cells. *J Immunol,* 182**,** 5419-5429.

Lu, J., Pei, H., Kaeck, M. and Thompson, H. J. 1997. Gene expression changes associated with chemically induced rat mammary carcinogenesis. *Mol Carcinog,* 20**,** 204-215.

Luo, J., Nikolaev, A. Y., Imai, S., Chen, D., Su, F., Shiloh, A., Guarente, L. and Gu, W. 2001. Negative control of p53 by Sir2α promotes cell survival under stress. *Cell,* 107**,** 137-148.

Luthi-Carter, R., Taylor, D. M., Pallos, J., Lambert, E., Amore, A., Parker, A., Moffitt, H., Smith, D. L., Runne, H., Gokce, O., Kuhn, A., Xiang, Z., Maxwell, M. M., Reeves, S. A., Bates, G. P., Neri, C., Thompson, L. M., Marsh, J. L. and Kazantsev, A. G. 2010. SIRT2 inhibition achieves neuroprotection by decreasing sterol biosynthesis. *Proc Natl Acad Sci U S A,* 107**,** 7927-7932.

Lyu, Y., K., Lim, K. R., Lee, B. Y., Kim, K. S. and Lee, W. Y. 2008. Microgravimetric lectin biosensor based on signal amplification using carbohydrate-stabilized gold nanoparticles. *Chem Commun (Camb)***,** 4771-4773.

Madsen, P., Rasmussen, H. H., Flint, T., Gromov, P., Kruse, T. A., Honore, B., Vorum, H. and Celis, J. E. 1995. Cloning, expression, and chromosome mapping of human galectin-7. *J Biol Chem,* 270**,** 5823-5829.

Magnaldo, T., Bernerd, F. and Darmon, M. 1995. Galectin-7, a human 14-kDa S-lectin, specifically expressed in keratinocytes and sensitive to retinoic acid. *Dev Biol,* 168**,** 259-271.

Magnaldo, T., Fowlis, D. and Darmon, M. 1998. Galectin-7, a marker of all types of stratified epithelia. *Differentiation,* 63**,** 159-168.

Mahlknecht, U., Ho, A. D. and Voelter-Mahlknecht, S. 2006. Chromosomal organization and fluorescence *in situ* hybridization of the human Sirtuin 6 gene. *Int J Oncol,* 28**,** 447-456.

Martinez-Pastor, B. and Mostoslavsky, R. 2012. Sirtuins, metabolism, and cancer. *Front Pharmacol,* 3**,** 22.

Masuyer, G., Jabeen, T., Oberg, C. T., Leffler, H., Nilsson, U. J. and Acharya, K. R. 2012. Inhibition mechanism of human galectin-7 by a novel galactose-benzylphosphate inhibitor. *FEBS J,* 279**,** 193-202.

Mccay, C. M., Crowell, M. F. and Maynard, L. A. 1935. The effect of retarded growth upon the length of life span and upon the ultimate body size. . *J Nutr,* 10**,** 63-79.

Mccoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. 2007. Phaser crystallographic software. *J App Crystallogr,* 40**,** 658-674.

Mcguinness, D., Mcguinness, D. H., Mccaul, J. A. and Shiels, P. G. 2011. Sirtuins, bioageing, and cancer. *J Aging Res,* 2011**,** 235754.

Menkhorst, E., Koga, K., Van Sinderen, M. and Dimitriadis, E. 2014a. Galectin-7 serum levels are altered prior to the onset of pre-eclampsia. *Placenta,* 35**,** 281-285.

Menkhorst, E. M., Gamage, T., Cuman, C., Kaitu'u-Lino, T. J., Tong, S. and Dimitriadis, E. 2014b. Galectin-7 acts as an adhesion molecule during implantation and increased expression is associated with miscarriage. *Placenta,* 35**,** 195-201.

Michan, S. and Sinclair, D. 2007. Sirtuins in mammals: insights into their biological function. *Biochem J,* 404**,** 1-13.

Michel, A. K., Nangia-Makker, P., Raz, A. and Cloninger, M. J. 2014. Lactose-Functionalized Dendrimers Arbitrate the Interaction of Galectin-3/MUC1 Mediated Cancer Cellular Aggregation. *Chembiochem,* 15**,** 2106-2112.

Michishita, E., Mccord, R. A., Boxer, L. D., Barber, M. F., Hong, T., Gozani, O. and Chua, K. F. 2009. Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human SIRT6. *Cell Cycle,* 8**,** 2664-2666.

Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C. and Horikawa, I. 2005. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. *Mol Biol Cell,* 16**,** 4623-4635.

Miller, S. I., Ernst, R. K. and Bader, M. W. 2005. LPS, TLR4 and infectious disease diversity. *Nat Rev Microbiol,* 3**,** 36-46.

Milne, J. C., Lambert, P. D., Schenk, S., Carney, D. P., Smith, J. J., Gagne, D. J., Jin, L., Boss, O., Perni, R. B., Vu, C. B., Bemis, J. E., Xie, R., Disch, J. S., Ng, P. Y., Nunes, J. J., Lynch, A. V., Yang, H., Galonek, H., Israelian, K., Choy, W., Iffland, A., Lavu, S., Medvedik, O., Sinclair, D. A., Olefsky, J. M., Jirousek, M. R., Elliott, P. J. and Westphal, C. H. 2007. Small molecule activators of SIRT1 as therapeutics for the treatment of type 2 diabetes. *Nature,* 450**,** 712-716.

Min, S. W., Cho, S. H., Zhou, Y., Schroeder, S., Haroutunian, V., Seeley, W. W., Huang, E. J., Shen, Y., Masliah, E., Mukherjee, C., Meyers, D., Cole, P. A., Ott, M. and Gan, L. 2010. Acetylation of tau inhibits its degradation and contributes to tauopathy. *Neuron,* 67**,** 953-966.

Miura, Y., Harumiya, S., Ono, K., Fujimoto, E., Akiyama, M., Fujii, N., Kawano, H., Wachi, H. and Tajima, S. 2013. Galectin-7 and actin are components of amyloid deposit of localized cutaneous amyloidosis. *Exp Dermatol,* 22**,** 36-40.

Moniot, S., Schutkowski, M. and Steegborn, C. 2013. Crystal structure analysis of human Sirt2 and its ADP-ribose complex. *J Struct Biol,* 182**,** 136-143.

Moniot, S., Weyand, M. and Steegborn, C. 2012. Structures, substrates, and regulators of Mammalian sirtuins - opportunities and challenges for drug development. *Front Pharmacol,* 3**,** 16.

Morris S, Ahmad N, Andre S, Kaltner H, Gabius Hj, Brenowitz M and Brewer F 2004. Quaternary solution structures of galectins-1, -3, and -7. *Glycobiology,* 14**,** 293-300.

Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M. M., Mills, K. D., Patel, P., Hsu, J. T., Hong, A. L., Ford, E., Cheng, H. L., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M. O., Hursting, S., Barrett, J. C., Guarente, L., Mulligan, R., Demple, B., Yancopoulos, G. D. and Alt, F. W. 2006. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell,* 124**,** 315-329.

Motta, M. C., Divecha, N., Lemieux, M., Kamel, C., Chen, D., Gu, W., Bultsma, Y., Mcburney, M. and Guarente, L. 2004. Mammalian SIRT1 represses forkhead transcription factors. *Cell,* 116**,** 551-563.

Mouline, Z., Mahon, E., Gomez, E., Barragan-Montero, V., Montero, J. L. and Barboiu, M. 2014. Entropy-driven lectin-recognition of multivalent glycovesicles. *Chem Commun (Camb),* 50**,** 731-733.

Munoz-Suano, A., Hamilton, A. B. and Betz, A. G. 2011. Gimme shelter: the immune system during pregnancy. *Immunol Rev,* 241**,** 20-38.

Muth, V., Nadaud, S., Grummt, I. and Voit, R. 2001. Acetylation of TAF(I)68, a subunit of TIF-IB/SL1, activates RNA polymerase I transcription. *EMBO J,* 20**,** 1353-1362.

Nagae, M. and Yamaguchi, Y. 2014. Three-dimensional structural aspects of protein-polysaccharide interactions. *Int J Mol Sci,* 15**,** 3768-3783.

Nakahata, Y., Kaluzova, M., Grimaldi, B., Sahar, S., Hirayama, J., Chen, D., Guarente, L. P. and Sassone-Corsi, P. 2008. The NAD^+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. *Cell,* 134**,** 329-340.

Nakamura, Y., Ogura, M., Tanaka, D. and Inagaki, N. 2008. Localization of mouse mitochondrial SIRT proteins: shift of SIRT3 to nucleus by co-expression with SIRT5. *Biochem Biophys Res Commun,* 366**,** 174-179.

Nesmelova, I. V., Berbis, M. A., Miller, M. C., Canada, F. J., Andre, S., Jimenez-Barbero, J., Gabius, H. J. and Mayo, K. H. 2012. 1H, 13C, and 15N backbone and side-chain chemical shift assignments for the 31 kDa human galectin-7 (p53-induced gene 1) homodimer, a pro-apoptotic lectin. *Biomol NMR Assign,* 6**,** 127-129.

Newkome, G. R., Moorefield, C. N. and Vögtle, F. 2001. *Dendrimers and Dendrons: Concepts, Syntheses, Applications,* Weinheim, Germany, Wiley-VCH.

Newman, J. C., He, W. and Verdin, E. 2012. Mitochondrial protein acylation and intermediary metabolism: regulation by sirtuins and implications for metabolic disease. *J Biol Chem,* 287**,** 42436-42443.

Nguyen, G. T., Gertz, M. and Steegborn, C. 2013a. Crystal structures of sirt3 complexes with 4'-bromo-resveratrol reveal binding sites and inhibition mechanism. *Chem Biol,* 20**,** 1375-1385.

Nguyen, G. T., Schaefer, S., Gertz, M., Weyand, M. and Steegborn, C. 2013b. Structures of human sirtuin 3 complexes with ADP-ribose and with carba-NAD^+ and SRT1720: binding details and inhibition mechanism. *Acta Crystallogr D Biol Crystallogr,* 69**,** 1423-1432.

Nickel, W. 2003. The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. *Eur J Biochem,* 270**,** 2109-2119.

Nieminen, J., St-Pierre, C., Bhaumik, P., Poirier, F. and Sato, S. 2008. Role of galectin-3 in leukocyte recruitment in a murine model of lung infection by Streptococcus pneumoniae. *J Immunol,* 180**,** 2466-2473.

Nio-Kobayashi, J., Takahashi-Iwanaga, H. and Iwanaga, T. 2009. Immunohistochemical localization of six galectin subtypes in the mouse digestive tract. *J Histochem Cytochem,* 57**,** 41-50.

North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M. and Verdin, E. 2003. The human Sir2 ortholog, SIRT2, is an NAD^+-dependent tubulin deacetylase. *Mol Cell,* 11**,** 437-444.

Nowell, P. C. 1960. Phytohemagglutinin: an initiator of mitosis in cultures of normal human leukocytes. *Cancer Res,* 20**,** 462-466.

O'hagan, H. M., Mohammad, H. P. and Baylin, S. B. 2008. Double strand breaks can initiate gene silencing and SIRT1-dependent onset of DNA methylation in an exogenous promoter CpG island. *PLoS Genet,* 4**,** e1000155.

Ochieng, J., Furtak, V. and Lukyanov, P. 2004. Extracellular functions of galectin-3. *Glycoconj J,* 19**,** 527-535.

Oka, T., Murakami, S., Arata, Y., Hirabayashi, J., Kasai, K., Wada, Y. and Futai, M. 1999. Identification and cloning of rat galectin-2: expression is predominantly in epithelial cells of the stomach. *Arch Biochem Biophys,* 361**,** 195-201.

Onyango, P., Celic, I., Mccaffery, J. M., Boeke, J. D. and Feinberg, A. P. 2002. SIRT3, a human SIR2 homologue, is an NAD-dependent deacetylase localized to mitochondria. *Proc Natl Acad Sci U S A,* 99**,** 13653-13658.

Ouaissi, M., Sielezneff, I., Silvestre, R., Sastre, B., Bernard, J. P., Lafontaine, J. S., Payan, M. J., Dahan, L., Pirro, N., Seitz, J. F., Mas, E., Lombardo, D. and Ouaissi, A. 2008. High histone deacetylase 7 (HDAC7) expression is significantly associated with adenocarcinomas of the pancreas. *Ann Surg Oncol,* 15**,** 2318-2328.

Pan, M., Yuan, H., Brent, M., Ding, E. C. and Marmorstein, R. 2012. SIRT1 contains N- and C-terminal regions that potentiate deacetylase activity. *J Biol Chem,* 287**,** 2468-2476.

Pan, P. W., Feldman, J. L., Devries, M. K., Dong, A., Edwards, A. M. and Denu, J. M. 2011. Structure and biochemical functions of SIRT6. *J Biol Chem,* 286**,** 14575-14587.

Paredes, S., Villanova, L. and Chua, K. F. 2014. Molecular pathways: emerging roles of mammalian Sirtuin SIRT7 in cancer. *Clin Cancer Res,* 20**,** 1741-1746.

Park, S. J., Ahmad, F., Philp, A., Baar, K., Williams, T., Luo, H., Ke, H., Rehmann, H., Taussig, R., Brown, A. L., Kim, M. K., Beaven, M. A., Burgin, A. B., Manganiello, V. and Chung, J. H. 2012. Resveratrol ameliorates aging-related metabolic phenotypes by inhibiting cAMP phosphodiesterases. *Cell,* 148**,** 421-433.

Peck, B., Chen, C. Y., Ho, K. K., Di Fruscia, P., Myatt, S. S., Coombes, R. C., Fuchter, M. J., Hsiao, C. D. and Lam, E. W. 2010. SIRT inhibitors induce cell death and p53 acetylation through targeting both SIRT1 and SIRT2. *Mol Cancer Ther,* 9**,** 844-855.

Percec, V., Leowanawat, P., Sun, H. J., Kulikov, O., Nusbaum, C. D., Tran, T. M., Bertin, A., Wilson, D. A., Peterca, M., Zhang, S., Kamat, N. P., Vargo, K., Moock, D., Johnston, E. D., Hammer, D. A., Pochan, D. J., Chen, Y., Chabre, Y. M., Shiao, T. C., Bergeron-Brlek, M., Andre, S., Roy, R., Gabius, H. J. and Heiney, P. A. 2013. Modular synthesis of amphiphilic Janus glycodendrimers and their self-assembly into glycodendrimersomes and other complex architectures with bioactivity to biomedically relevant lectins. *J Am Chem Soc,* 135**,** 9055-9077.

Pfluger, P. T., Herranz, D., Velasco-Miguel, S., Serrano, M. and Tschop, M. H. 2008. Sirt1 protects against high-fat diet-induced metabolic damage. *Proc Natl Acad Sci U S A,* 105**,** 9793-9798.

Pillai, J. B., Isbatan, A., Imai, S. and Gupta, M. P. 2005. Poly(ADP-ribose) polymerase-1-dependent cardiac myocyte cell death during heart failure is mediated by NAD^+ depletion and reduced Sir2α deacetylase activity. *J Biol Chem,* 280**,** 43121-43130.

Pillai, V. B., Sundaresan, N. R., Kim, G., Gupta, M., Rajamohan, S. B., Pillai, J. B., Samant, S., Ravindra, P. V., Isbatan, A. and Gupta, M. P. 2010. Exogenous NAD blocks cardiac hypertrophic response via activation of the SIRT3-LKB1-AMP-activated kinase pathway. *J Biol Chem,* 285**,** 3133-3144.

Planavila, A., Iglesias, R., Giralt, M. and Villarroya, F. 2011. Sirt1 acts in association with PPARalpha to protect the heart from hypertrophy, metabolic dysregulation, and inflammation. *Cardiovasc Res,* 90**,** 276-284.

Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W. and Vogelstein, B. 1997. A model for p53-induced apoptosis. *Nature,* 389**,** 300-305.

Ponugoti, B., Kim, D. H., Xiao, Z., Smith, Z., Miao, J., Zang, M., Wu, S. Y., Chiang, C. M., Veenstra, T. D. and Kemper, J. K. 2010. SIRT1 deacetylates and inhibits SREBP-1C activity in regulation of hepatic lipid metabolism. *J Biol Chem,* 285**,** 33959-33970.

Pruitt, K., Zinn, R. L., Ohm, J. E., Mcgarvey, K. M., Kang, S. H., Watkins, D. N., Herman, J. G. and Baylin, S. B. 2006. Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. *PLoS Genet,* 2**,** e40.

Qiang, L., Lin, H. V., Kim-Muller, J. Y., Welch, C. L., Gu, W. and Accili, D. 2011. Proatherogenic abnormalities of lipid metabolism in SirT1 transgenic mice are mediated through Creb deacetylation. *Cell Metab,* 14**,** 758-767.

Qiu, X., Brown, K., Hirschey, M. D., Verdin, E. and Chen, D. 2010. Calorie restriction reduces oxidative stress by SIRT3-mediated SOD2 activation. *Cell Metab,* 12**,** 662-667.

Rabinovich, G. A. 1999. Galectins: an evolutionarily conserved family of animal lectins with multifunctional properties; a trip from the gene to clinical therapy. *Cell Death Differ,* 6**,** 711-721.

Rabinovich, G. A. and Croci, D. O. 2012. Regulatory circuits mediated by lectin-glycan interactions in autoimmunity and cancer. *Immunity,* 36**,** 322-335.

Rabinovich, G. A., Daly, G., Dreja, H., Tailor, H., Riera, C. M., Hirabayashi, J. and Chernajovsky, Y. 1999. Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis. *J Exp Med,* 190**,** 385-398.

Rabinovich, G. A., Rubinstein, N. and Toscano, M. A. 2002. Role of galectins in inflammatory and immunomodulatory processes. *Biochim Biophys Acta,* 1572**,** 274-284.

Rabinovich, G. A. and Toscano, M. A. 2009. Turning 'sweet' on immunity: galectin-glycan interactions in immune tolerance and inflammation. *Nat Rev Immunol,* 9**,** 338-352.

Rabinovich, G. A., Toscano, M. A., Jackson, S. S. and Vasta, G. R. 2007. Functions of cell surface galectin-glycoprotein lattices. *Curr Opin Struct Biol,* 17**,** 513-520.

Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. 1963. Stereochemistry of polypeptide chain configurations. *J Mol Biol,* 7**,** 95-99.

Reuter, G. and Gabius, H. J. 1999. Eukaryotic glycosylation: whim of nature or multipurpose tool? *Cell Mol Life Sci,* 55**,** 368-422.

Reynolds, A. J., Haines, A. H. and Russell, D. A. 2006. Gold glyconanoparticles for mimics and measurement of metal ion-mediated carbohydrate-carbohydrate interactions. *Langmuir,* 22**,** 1156-1163.

Rine, J. and Herskowitz, I. 1987. Four genes responsible for a position effect on expression from HML and HMR in *Saccharomyces cerevisiae*. *Genetics,* 116**,** 9-22.

Rodgers, J. T., Lerin, C., Haas, W., Gygi, S. P., Spiegelman, B. M. and Puigserver, P. 2005. Nutrient control of glucose homeostasis through a complex of PGC-1α and SIRT1. *Nature,* 434**,** 113-118.

Rogina, B. and Helfand, S. L. 2004. Sir2 mediates longevity in the fly through a pathway related to calorie restriction. *Proc Natl Acad Sci U S A,* 101**,** 15998-16003.

Rossmann, M. G. and Blow, D. M. 1962. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr,* 15**,** 24-31.

Roth, M. and Chen, W. Y. 2014. Sorting out functions of sirtuins in cancer. *Oncogene,* 33**,** 1609-1620.

Rothgiesser, K. M., Erener, S., Waibel, S., Luscher, B. and Hottiger, M. O. 2010. SIRT2 regulates NF-kappaB dependent gene expression through deacetylation of p65 Lys310. *J Cell Sci,* 123**,** 4251-4258.

Roy, R. and Kim, J. M. 1999. Amphiphilic p-tert-Butylcalix[4]arene Scaffolds Containing Exposed Carbohydrate Dendrons. *Angewandte Chemie International Edition,* 38**,** 369-372.

Roy, R., Zanini, D., Meunier, S. J. and Romanowska, A. 1993. Solid-phase synthesis of dendritic sialoside inhibitors of influenza A virus haemagglutinin. *Chem Commun (Camb)***,** 1869-1872.

Rupp, B. 2010. *Biomolecular crystallography,* Garland Science New York.

Russo Krauss, I., Merlino, A., Vergara, A. and Sica, F. 2013. An overview of biological macromolecule crystallization. *Int J Mol Sci,* 14**,** 11643-11691.

Sacchettini, J. C., Baum, L. G. and Brewer, C. F. 2001. Multivalent protein-carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction. *Biochemistry,* 40**,** 3009-3015.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* 239**,** 487-491.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science,* 230**,** 1350-1354.

Salvatore, P., Benvenuto, G., Pero, R., Lembo, F., Bruni, C. B. and Chiariotti, L. 2000. Galectin-1 gene expression and methylation state in human T leukemia cell lines. *Int J Oncol,* 17**,** 1015-1018.

Sanders, B. D., Jackson, B. and Marmorstein, R. 2010. Structural basis for sirtuin function: what we know and what we don't. *Biochim Biophys Acta,* 1804**,** 1604-1616.

Santucci, L., Fiorucci, S., Rubinstein, N., Mencarelli, A., Palazzetti, B., Federici, B., Rabinovich, G. A. and Morelli, A. 2003. Galectin-1 suppresses experimental colitis in mice. *Gastroenterology,* 124**,** 1381-1394.

Saraboji, K., Hakansson, M., Genheden, S., Diehl, C., Qvist, J., Weininger, U., Nilsson, U. J., Leffler, H., Ryde, U., Akke, M. and Logan, D. T. 2012. The carbohydrate-binding site in galectin-3 is preorganized to recognize a sugarlike framework of oxygens: ultra-high-resolution structures and water dynamics. *Biochemistry,* 51**,** 296-306.

Sato, M., Nishi, N., Shoji, H., Kumagai, M., Imaizumi, T., Hata, Y., Hirashima, M., Suzuki, S. and Nakamura, T. 2002. Quantification of galectin-7 and its localization in adult mouse tissues. *J Biochem,* 131**,** 255-260.

Sato, S., Burdett, I. and Hughes, R. C. 1993. Secretion of the baby hamster kidney 30-kDa galactose-binding lectin from polarized and nonpolarized cells: a pathway independent of the endoplasmic reticulum-Golgi complex. *Exp Cell Res,* 207**,** 8-18.

Satoh, A., Brace, C. S., Rensing, N., Cliften, P., Wozniak, D. F., Herzog, E. D., Yamada, K. A. and Imai, S. 2013. Sirt1 extends life span and delays aging in mice through the regulation of Nk2 homeobox 1 in the DMH and LH. *Cell Metab,* 18**,** 416-430.

Saunders, L. R. and Verdin, E. 2007. Sirtuins: critical regulators at the crossroads between cancer and aging. *Oncogene,* 26**,** 5489-5504.

Saussez, S. and Kiss, R. 2006. Galectin-7. *Cell Mol Life Sci,* 63**,** 686-697.

Sauve, A. A. 2010. Sirtuin chemical mechanisms. *Biochim Biophys Acta,* 1804**,** 1591-1603.

Scher, M. B., Vaquero, A. and Reinberg, D. 2007. SirT3 is a nuclear NAD^+-dependent histone deacetylase that translocates to the mitochondria upon cellular stress. *Genes Dev,* 21**,** 920-928.

Schmeisser, K., Mansfeld, J., Kuhlow, D., Weimer, S., Priebe, S., Heiland, I., Birringer, M., Groth, M., Segref, A., Kanfi, Y., Price, N. L., Schmeisser, S., Schuster, S., Pfeiffer, A. F., Guthke, R., Platzer, M., Hoppe, T., Cohen, H. Y., Zarse, K.,

Sinclair, D. A. and Ristow, M. 2013. Role of sirtuins in lifespan regulation is linked to methylation of nicotinamide. *Nat Chem Biol,* 9**,** 693-700.

Schuetz, A., Min, J., Antoshenko, T., Wang, C. L., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R. and Plotnikov, A. N. 2007. Structural basis of inhibition of the human NAD^+-dependent deacetylase SIRT5 by suramin. *Structure,* 15**,** 377-389.

Schwer, B., North, B. J., Frye, R. A., Ott, M. and Verdin, E. 2002. The human silent information regulator (Sir)2 homologue hSIRT3 is a mitochondrial nicotinamide adenine dinucleotide-dependent deacetylase. *J Cell Biol,* 158**,** 647-657.

Seligson, D. B., Horvath, S., Mcbrian, M. A., Mah, V., Yu, H., Tze, S., Wang, Q., Chia, D., Goodglick, L. and Kurdistani, S. K. 2009. Global levels of histone modifications predict prognosis in different cancers. *Am J Pathol,* 174**,** 1619-1628.

Sharon, N. and Lis, H. 1972. Lectins: cell-agglutinating and sugar-specific proteins. *Science,* 177**,** 949-959.

Sharon, N. and Lis, H. 2004. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology,* 14**,** 53r-62r.

Shiao, T. C. and Roy, R. 2012. Glycodendrimers as functional antigens and antitumor vaccines. *New J Chem,* 36**,** 324-339.

Shimazu, T., Hirschey, M. D., Hua, L., Dittenhafer-Reed, K. E., Schwer, B., Lombard, D. B., Li, Y., Bunkenborg, J., Alt, F. W., Denu, J. M., Jacobson, M. P. and Verdin, E. 2010. SIRT3 deacetylates mitochondrial 3-hydroxy-3-methylglutaryl CoA synthase 2 and regulates ketone body production. *Cell Metab,* 12**,** 654-661.

Shin, J., He, M., Liu, Y., Paredes, S., Villanova, L., Brown, K., Qiu, X., Nabavi, N., Mohrin, M., Wojnoonski, K., Li, P., Cheng, H. L., Murphy, A. J., Valenzuela, D. M., Luo, H., Kapahi, P., Krauss, R., Mostoslavsky, R., Yancopoulos, G. D., Alt, F. W., Chua, K. F. and Chen, D. 2013. SIRT7 represses Myc activity to suppress ER stress and prevent fatty liver disease. *Cell Rep,* 5**,** 654-665.

Shore, D., Squire, M. and Nasmyth, K. A. 1984. Characterization of two genes required for the position-effect control of yeast mating-type genes. *EMBO J,* 3**,** 2817-2823.

Sigal, A. and Rotter, V. 2000. Oncogenic mutations of the p53 tumor suppressor: the demons of the guardian of the genome. *Cancer Res,* 60**,** 6788-6793.

Sinclair, D. A. and Guarente, L. 2014. Small-molecule allosteric activators of sirtuins. *Annu Rev Pharmacol Toxicol,* 54**,** 363-380.

Smith, B. C. and Denu, J. M. 2007. Acetyl-lysine analog peptides as mechanistic probes of protein deacetylases. *J Biol Chem,* 282**,** 37256-37265.

Smith, J. S. and Boeke, J. D. 1997. An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev,* 11**,** 241-254.

Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C. and Boeke, J. D. 2000. A phylogenetically conserved NAD^+-dependent protein deacetylase activity in the Sir2 protein family. *Proc Natl Acad Sci U S A,* 97**,** 6658-6663.

Smyth, D. R., Mrozkiewicz, M. K., Mcgrath, W. J., Listwan, P. and Kobe, B. 2003. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci,* 12**,** 1313-1322.

Someya, S., Yu, W., Hallows, W. C., Xu, J., Vann, J. M., Leeuwenburgh, C., Tanokura, M., Denu, J. M. and Prolla, T. A. 2010. Sirt3 mediates reduction of oxidative damage and prevention of age-related hearing loss under caloric restriction. *Cell,* 143**,** 802-812.

St-Pierre, Y., Campion, C. G. and Grosset, A. A. 2012. A distinctive role for galectin-7 in cancer ? *Front Biosci (Landmark Ed),* 17**,** 438-450.

Stannard, K. A., Collins, P. M., Ito, K., Sullivan, E. M., Scott, S. A., Gabutero, E., Darren Grice, I., Low, P., Nilsson, U. J., Leffler, H., Blanchard, H. and Ralph, S. J. 2010. Galectin inhibitory disaccharides promote tumour immunity in a breast cancer model. *Cancer Lett,* 299**,** 95-110.

Stillman, B. N., Hsu, D. K., Pang, M., Brewer, C. F., Johnson, P., Liu, F. T. and Baum, L. G. 2006. Galectin-3 and galectin-1 bind distinct cell surface glycoprotein receptors to induce T cell death. *J Immunol,* 176**,** 778-789.

Stockert, R. J., Morell, A. G. and Scheinberg, I. H. 1974. Mammalian hepatic lectin. *Science,* 186**,** 365-366.

Stowell, S. R., Dias-Baruffi, M., Penttila, L., Renkonen, O., Nyame, A. K. and Cummings, R. D. 2004. Human galectin-1 recognition of poly-N-acetyllactosamine and chimeric polysaccharides. *Glycobiology,* 14**,** 157-167.

Stunkel, W., Peh, B. K., Tan, Y. C., Nayagam, V. M., Wang, X., Salto-Tellez, M., Ni, B., Entzeroth, M. and Wood, J. 2007. Function of the SIRT1 protein deacetylase in cancer. *Biotechnol J,* 2**,** 1360-1368.

Sturm, A., Lensch, M., Andre, S., Kaltner, H., Wiedenmann, B., Rosewicz, S., Dignass, A. U. and Gabius, H. J. 2004. Human galectin-2: novel inducer of T cell apoptosis with distinct profile of caspase activation. *J Immunol,* 173**,** 3825-3837.

Sundaresan, N. R., Gupta, M., Kim, G., Rajamohan, S. B., Isbatan, A. and Gupta, M. P. 2009. Sirt3 blocks the cardiac hypertrophic response by augmenting Foxo3a-dependent antioxidant defense mechanisms in mice. *J Clin Invest,* 119**,** 2758-2771.

Sundaresan, N. R., Vasudevan, P., Zhong, L., Kim, G., Samant, S., Parekh, V., Pillai, V. B., Ravindra, P. V., Gupta, M., Jeevanandam, V., Cunningham, J. M., Deng, C. X., Lombard, D. B., Mostoslavsky, R. and Gupta, M. P. 2012. The sirtuin SIRT6 blocks IGF-Akt signaling and development of cardiac hypertrophy by targeting c-Jun. *Nat Med,* 18**,** 1643-1650.

Szczepankiewicz, B. G., Dai, H., Koppetsch, K. J., Qian, D., Jiang, F., Mao, C. and Perni, R. B. 2012. Synthesis of carba-NAD and the structures of its ternary complexes with SIRT3 and SIRT5. *J Org Chem,* 77**,** 7319-7329.

Tanner, K. G., Landry, J., Sternglanz, R. and Denu, J. M. 2000. Silent information regulator 2 family of NAD- dependent histone/protein deacetylases generates a unique product, 1-O-acetyl-ADP-ribose. *Proc Natl Acad Sci U S A,* 97**,** 14178-14182.

Tanno, M., Sakamoto, J., Miura, T., Shimamoto, K. and Horio, Y. 2007. Nucleocytoplasmic shuttling of the NAD^+-dependent histone deacetylase SIRT1. *J Biol Chem,* 282**,** 6823-6832.

Tanny, J. C., Dowd, G. J., Huang, J., Hilz, H. and Moazed, D. 1999. An enzymatic activity in the yeast Sir2 protein that is essential for gene silencing. *Cell,* 99**,** 735-745.

Tanny, J. C. and Moazed, D. 2001. Coupling of histone deacetylation to NAD breakdown by the yeast silencing protein Sir2: Evidence for acetyl transfer from substrate to an NAD breakdown product. *Proc Natl Acad Sci U S A,* 98**,** 415-420.

Teichberg, V. I., Silman, I., Beitsch, D. D. and Resheff, G. 1975. A β-D-galactoside binding protein from electric organ tissue of Electrophorus electricus. *Proc Natl Acad Sci U S A,* 72**,** 1383-1387.

Tennen, R. I., Berber, E. and Chua, K. F. 2010. Functional dissection of SIRT6: identification of domains that regulate histone deacetylase activity and chromatin localization. *Mech Ageing Dev,* 131**,** 185-192.

Teo, I., Toms, S. M., Marteyn, B., Barata, T. S., Simpson, P., Johnston, K. A., Schnupf, P., Puhar, A., Bell, T., Tang, C., Zloh, M., Matthews, S., Rendle, P. M., Sansonetti, P. J. and Shaunak, S. 2012. Preventing acute gut wall damage in infectious diarrhoeas with glycosylated dendrimers. *EMBO Mol Med,* 4**,** 866-881.

Thompson, J. S., Ling, X. and Grunstein, M. 1994. Histone H3 amino terminus is required for telomeric and silent mating locus repression in yeast. *Nature,* 369**,** 245-247.

Thurston, T. L., Wandel, M. P., Von Muhlinen, N., Foeglein, A. and Randow, F. 2012. Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion. *Nature,* 482**,** 414-418.

Timmons, P. M., Colnot, C., Cail, I., Poirier, F. and Magnaldo, T. 1999. Expression of galectin-7 during epithelial development coincides with the onset of stratification. *Int J Dev Biol,* 43**,** 229-235.

Tissenbaum, H. A. and Guarente, L. 2001. Increased dosage of a sir-2 gene extends lifespan in *Caenorhabditis elegans*. *Nature,* 410**,** 227-230.

Tong, L. and Denu, J. M. 2010. Function and metabolism of sirtuin metabolite O-acetyl-ADP-ribose. *Biochim Biophys Acta,* 1804**,** 1617-1625.

Touaibia, M. and Roy, R. 2007. Glycodendrimers as anti-adhesion drugs against type 1 fimbriated *E. coli* uropathogenic infections. *Mini Rev Med Chem,* 7**,** 1270-1283.

Trapp, J., Jochum, A., Meier, R., Saunders, L., Marshall, B., Kunick, C., Verdin, E., Goekjian, P., Sippl, W. and Jung, M. 2006. Adenosine mimetics as inhibitors of NAD^+-dependent histone deacetylases, from kinase to sirtuin inhibition. *J Med Chem,* 49**,** 7307-7316.

Tsai, C. J., Sulman, E. P., Eifel, P. J., Jhingran, A., Allen, P. K., Deavers, M. T. and Klopp, A. H. 2013. Galectin-7 levels predict radiation response in squamous cell carcinoma of the cervix. *Gynecol Oncol,* 131**,** 645-649.

Tsang, A. W. and Escalante-Semerena, J. C. 1998. CobB, a new member of the SIR2 family of eucaryotic regulatory proteins, is required to compensate for the lack of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase activity in cobT mutants during cobalamin biosynthesis in *Salmonella typhimurium* LT2. *J Biol Chem,* 273**,** 31788-31794.

Tseng, A. H., Shieh, S. S. and Wang, D. L. 2013. SIRT3 deacetylates FOXO3 to protect mitochondria against oxidative damage. *Free Radic Biol Med,* 63**,** 222-234.

Tsuchiyama, Y., Wada, J., Zhang, H., Morita, Y., Hiragushi, K., Hida, K., Shikata, K., Yamamura, M., Kanwar, Y. S. and Makino, H. 2000. Efficacy of galectins in the amelioration of nephrotoxic serum nephritis in Wistar Kyoto rats. *Kidney Int,* 58**,** 1941-1952.

Uniprot 2014. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res,* 42**,** D191-198.

Vakhrusheva, O., Smolka, C., Gajawada, P., Kostin, S., Boettger, T., Kubin, T., Braun, T. and Bober, E. 2008. Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice. *Circ Res,* 102**,** 703-710.

Van Der Laan, A. M., Schirmer, S. H., De Vries, M. R., Koning, J. J., Volger, O. L., Fledderus, J. O., Bastiaansen, A. J., Hollander, M. R., Baggen, J. M., Koch, K. T., Baan, J., Jr., Henriques, J. P., Van Der Schaaf, R. J., Vis, M. M., Mebius, R. E., Van Der Pouw Kraan, T. C., Quax, P. H., Piek, J. J., Horrevoets, A. J. and Van Royen, N. 2012. Galectin-2 expression is dependent on the rs7291467 polymorphism and acts as an inhibitor of arteriogenesis. *Eur Heart J,* 33**,** 1076-1084.

Vaquero, A., Scher, M. B., Lee, D. H., Sutton, A., Cheng, H. L., Alt, F. W., Serrano, L., Sternglanz, R. and Reinberg, D. 2006. SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis. *Genes Dev,* 20**,** 1256-1261.

Vaquero, A., Sternglanz, R. and Reinberg, D. 2007. NAD+-dependent deacetylation of H4 lysine 16 by class III HDACs. *Oncogene,* 26**,** 5505-5520.

Varki, A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology,* 3**,** 97-130.

Vasta, G. R. 2009. Roles of galectins in infection. *Nat Rev Microbiol,* 7**,** 424-438.

Vaziri, H., Dessain, S. K., Ng Eaton, E., Imai, S. I., Frye, R. A., Pandita, T. K., Guarente, L. and Weinberg, R. A. 2001. hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. *Cell,* 107**,** 149-159.

Villalonga-Barber, C., Micha-Screttas, M., Steele, B. R., Georgopoulos, A. and Demetzos, C. 2008. Dendrimers as biopharmaceuticals: synthesis and properties. *Curr Top Med Chem,* 8**,** 1294-1309.

Viswanathan, M. and Guarente, L. 2011. Regulation of *Caenorhabditis elegans* lifespan by sir-2.1 transgenes. *Nature,* 477**,** E1-2.

Vögtle, F., Gabriele, R. and Werner, N. 2009. *Dendrimer Chemistry: Concepts, Syntheses, Properties, Applications,* Weinheim, Germany, Wiley-VCH.

Walker, A. K., Yang, F., Jiang, K., Ji, J. Y., Watts, J. L., Purushotham, A., Boss, O., Hirsch, M. L., Ribich, S., Smith, J. J., Israelian, K., Westphal, C. H., Rodgers, J. T., Shioda, T., Elson, S. L., Mulligan, P., Najafi-Shoushtari, H., Black, J. C., Thakur, J. K., Kadyk, L. C., Whetstine, J. R., Mostoslavsky, R., Puigserver, P., Li, X., Dyson, N. J., Hart, A. C. and Naar, A. M. 2010. Conserved role of SIRT1 orthologs in fasting-dependent inhibition of the lipid/cholesterol regulator SREBP. *Genes Dev,* 24**,** 1403-1417.

Wang, C., Chen, L., Hou, X., Li, Z., Kabra, N., Ma, Y., Nemoto, S., Finkel, T., Gu, W., Cress, W. D. and Chen, J. 2006. Interactions between E2F1 and SirT1 regulate apoptotic response to DNA damage. *Nat Cell Biol,* 8**,** 1025-1031.

Wang, F. and Tong, Q. 2009. SIRT2 suppresses adipocyte differentiation by deacetylating FOXO1 and enhancing FOXO1's repressive interaction with PPARgamma. *Mol Biol Cell,* 20**,** 801-808.

Wang, R. H., Sengupta, K., Li, C., Kim, H. S., Cao, L., Xiao, C., Kim, S., Xu, X., Zheng, Y., Chilton, B., Jia, R., Zheng, Z. M., Appella, E., Wang, X. W., Ried, T. and Deng, C. X. 2008. Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell,* 14**,** 312-323.

Wang, Y. and Tissenbaum, H. A. 2006. Overlapping and distinct functions for a Caenorhabditis elegans SIR2 and DAF-16/FOXO. *Mech Ageing Dev,* 127**,** 48-56.

Watkins, W. M. and Morgan, W. T. 1952. Neutralization of the anti-H agglutinin in eel serum by simple sugars. *Nature,* 169**,** 825-826.

Weiss, E. P. and Fontana, L. 2011. Caloric restriction: powerful protection for the aging heart and vasculature. *Am J Physiol Heart Circ Physiol,* 301**,** H1205-1219.

Wiersma, V. R., De Bruyn, M., Helfrich, W. and Bremer, E. 2013. Therapeutic potential of Galectin-9 in human disease. *Med Res Rev,* 33 Suppl 1**,** E102-126.

Wilson, J. M., Le, V. Q., Zimmerman, C., Marmorstein, R. and Pillus, L. 2006. Nuclear export modulates the cytoplasmic Sir2 homologue Hst2. *EMBO Rep,* 7**,** 1247-1251.

Winter, G. 2010. xia2: an expert system for macromolecular crystallography data reduction. *J App Crystallogr,* 43**,** 186-190.

Witt, O., Deubzer, H. E., Milde, T. and Oehme, I. 2009. HDAC family: What are the cancer relevant targets? *Cancer Lett,* 277**,** 8-21.

Wolfenden, M. L. and Cloninger, M. J. 2005. Mannose/glucose-functionalized dendrimers to investigate the predictable tunability of multivalent interactions. *J Am Chem Soc,* 127**,** 12168-12169.

Wu, J., Zhang, D., Chen, L., Li, J., Wang, J., Ning, C., Yu, N., Zhao, F., Chen, D., Chen, X., Chen, K., Jiang, H., Liu, H. and Liu, D. 2013. Discovery and mechanism study of SIRT1 activators that promote the deacetylation of fluorophore-labeled substrate. *J Med Chem,* 56**,** 761-780.

Wu, Z., Liu, M. C., Liang, M. and Fu, J. 2012. Sirt1 protects against thrombomodulin down-regulation and lung coagulation after particulate matter exposure. *Blood,* 119**,** 2422-2429.

Yamagata, K., Goto, Y., Nishimasu, H., Morimoto, J., Ishitani, R., Dohmae, N., Takeda, N., Nagai, R., Komuro, I., Suga, H. and Nureki, O. 2014. Structural basis for potent inhibition of SIRT2 deacetylase by a macrocyclic peptide inducing dynamic structural change. *Structure,* 22**,** 345-352.

Yoshida, H., Teraoka, M., Nishi, N., Nakakita, S., Nakamura, T., Hirashima, M. and Kamitori, S. 2010. X-ray structures of human galectin-9 C-terminal domain in complexes with a biantennary oligosaccharide and sialyllactose. *J Biol Chem,* 285**,** 36969-36976.

Yoshizaki, T., Milne, J. C., Imamura, T., Schenk, S., Sonoda, N., Babendure, J. L., Lu, J. C., Smith, J. J., Jirousek, M. R. and Olefsky, J. M. 2009. SIRT1 exerts anti-

inflammatory effects and improves insulin sensitivity in adipocytes. *Mol Cell Biol,* 29**,** 1363-1374.

Yu, S. S., Cai, Y., Ye, J. T., Pi, R. B., Chen, S. R., Liu, P. Q., Shen, X. Y. and Ji, Y. 2013. Sirtuin 6 protects cardiomyocytes from hypertrophy *in vitro* via inhibition of NF-κB-dependent transcriptional activity. *Br J Pharmacol,* 168**,** 117-128.

Zhao, K., Chai, X., Clements, A. and Marmorstein, R. 2003. Structure and autoregulation of the yeast Hst2 homolog of Sir2. *Nat Struct Biol,* 10**,** 864-871.

Zhao, W., Kruse, J. P., Tang, Y., Jung, S. Y., Qin, J. and Gu, W. 2008. Negative regulation of the deacetylase SIRT1 by DBC1. *Nature,* 451**,** 587-590.

Zhao, X., Allison, D., Condon, B., Zhang, F., Gheyi, T., Zhang, A., Ashok, S., Russell, M., Macewan, I., Qian, Y., Jamison, J. A. and Luz, J. G. 2013. The 2.5 A crystal structure of the SIRT1 catalytic domain bound to nicotinamide adenine dinucleotide (NAD^+) and an indole (EX527 analogue) reveals a novel mechanism of histone deacetylase inhibition. *J Med Chem,* 56**,** 963-969.

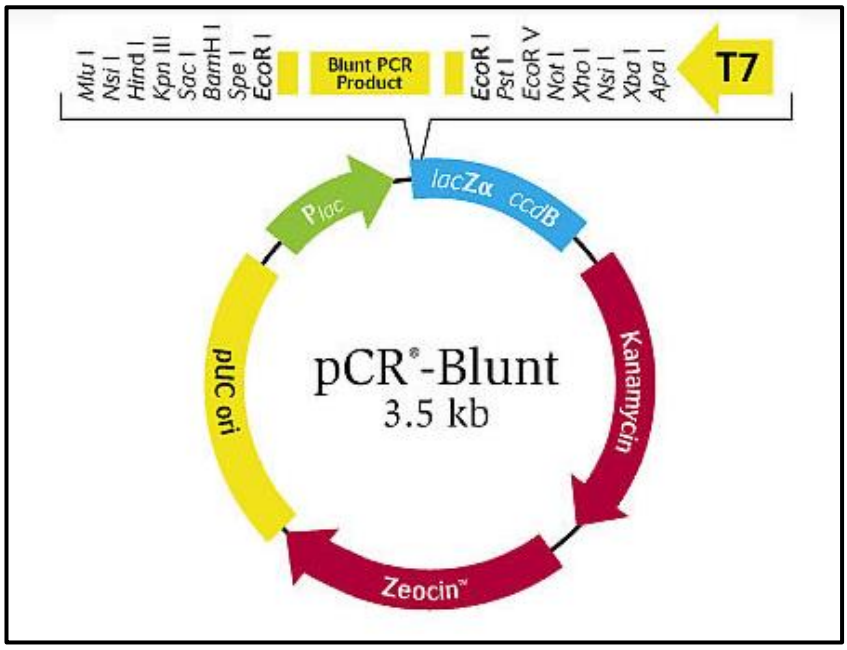
Zhong, L., D'urso, A., Toiber, D., Sebastian, C., Henry, R. E., Vadysirisack, D. D., Guimaraes, A., Marinelli, B., Wikstrom, J. D., Nir, T., Clish, C. B., Vaitheesvaran, B., Iliopoulos, O., Kurland, I., Dor, Y., Weissleder, R., Shirihai, O. S., Ellisen, L. W., Espinosa, J. M. and Mostoslavsky, R. 2010. The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1α. *Cell,* 140**,** 280-293.

Zhou, Y., Zhang, H., He, B., Du, J., Lin, H., Cerione, R. A. and Hao, Q. 2012. The bicyclic intermediate structure provides insights into the desuccinylation mechanism of human sirtuin 5 (SIRT5). *J Biol Chem,* 287**,** 28307-28314.

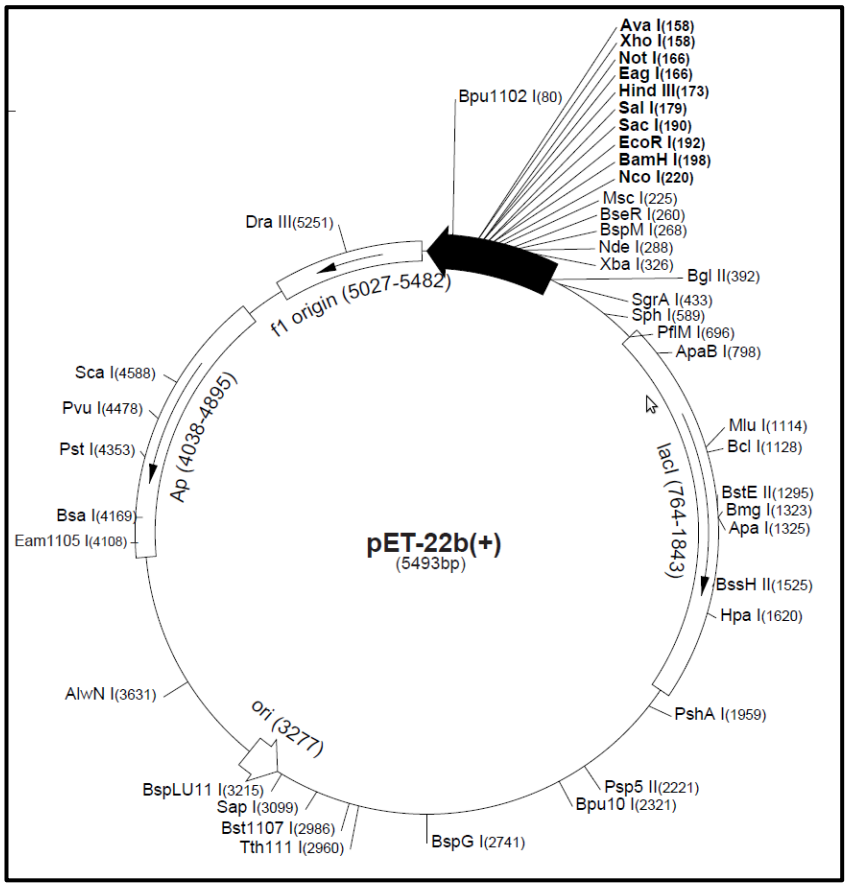
APPENDICES

APPENDIX A: Vector maps

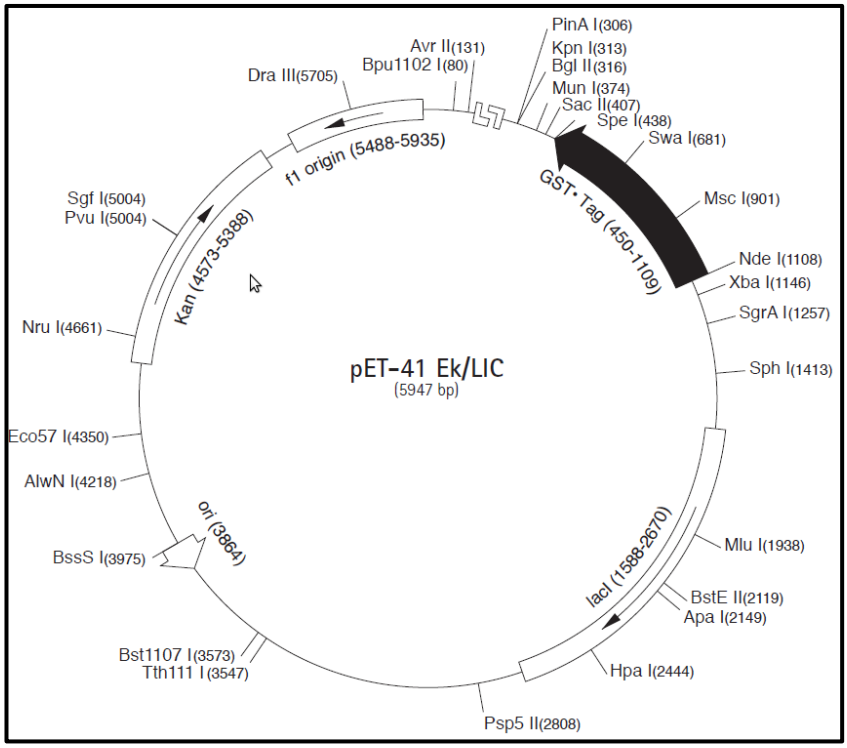
- **PCR blunt vector**



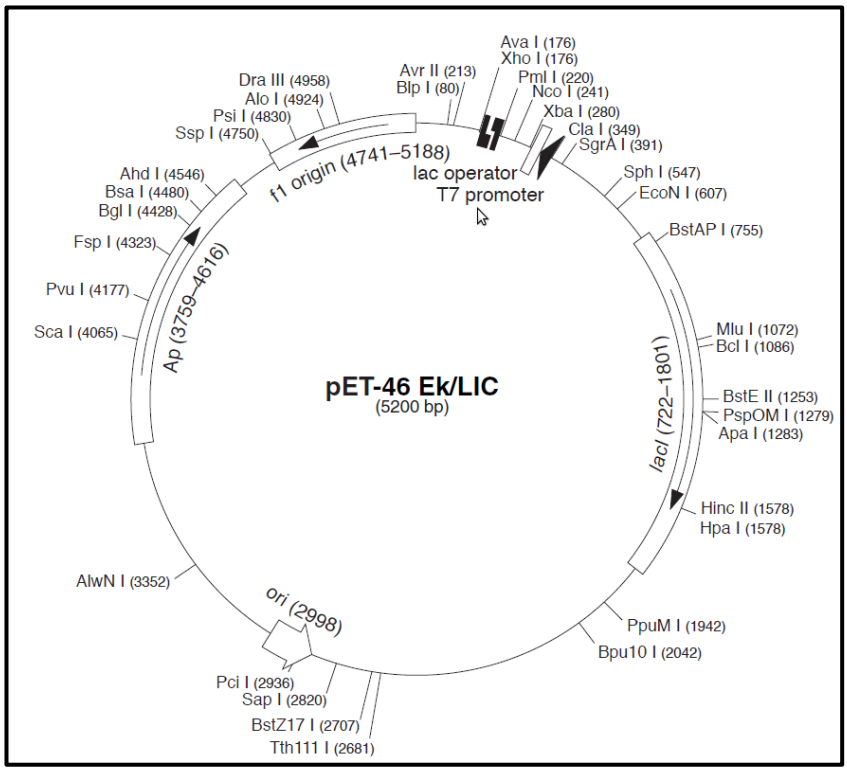
- **pET-22b vector**



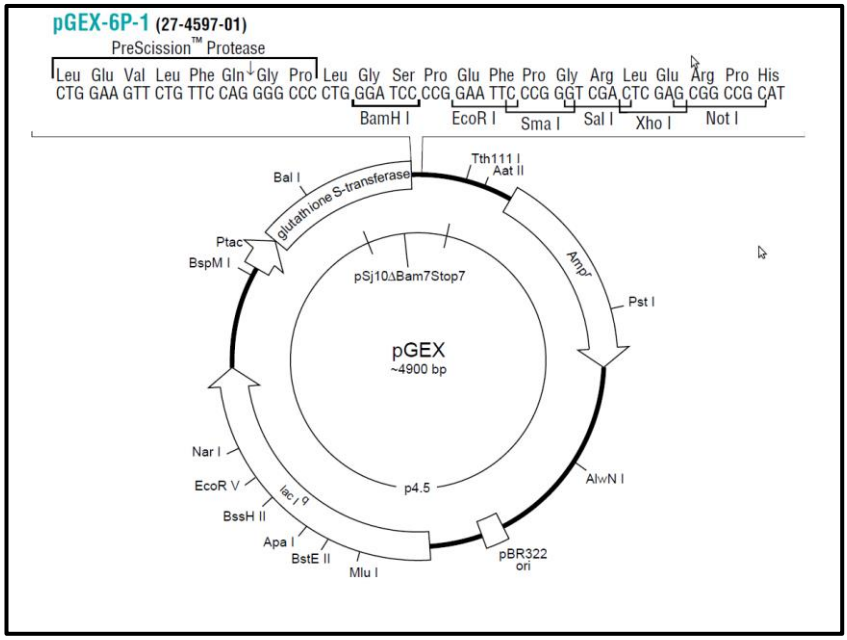
- **pET-41 Ek/LIC**



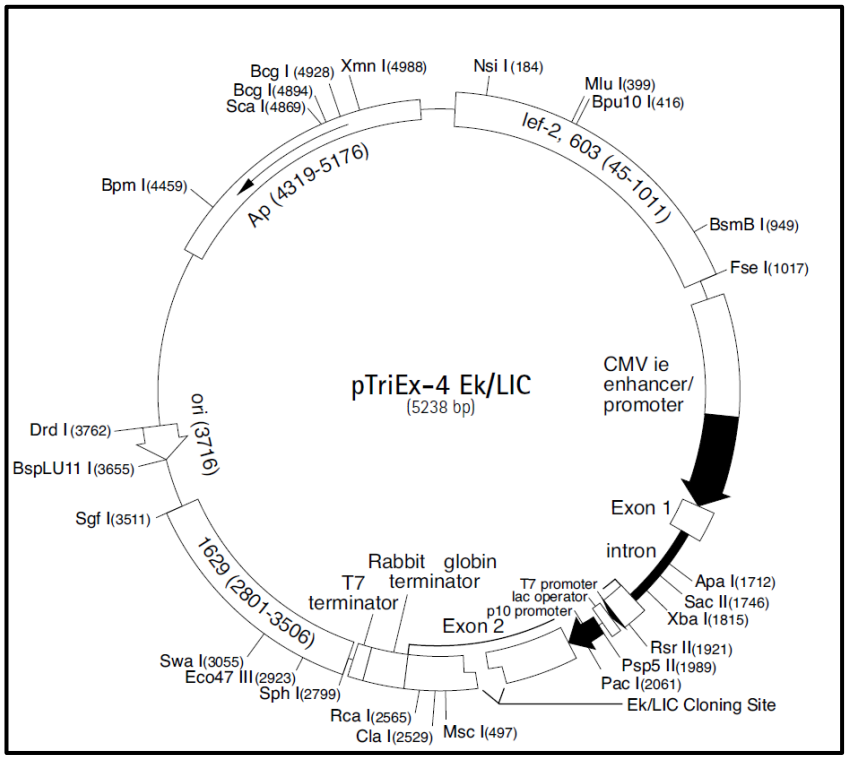
- **pET-46 Ek/LIC vector**



- **pGEx-6P-1 vector**



- **pTrieEx-4 Ek/LIC vector**



APPENDIX B: Composition of bacterial cell growth media

LB Media (q.s. 1000 mL)	
5 g	Yeast extract
10 g	Tryptone
10 g	Sodium Chloride

TB Media (q.s. 1000 mL)	
24 g	Yeast extract
12 g	Tryptone
4 mL	Glycerol
(q.s.)	900 mL
TB Salts (q.s. 100 mL)	
2.31 g	Potassium phosphate (monobasic)
12.54 g	Potassium phosphate (dibasic)

HIS Media (q.s. 1000 mL)	
50 g	Yeast extract
10 g	Casamino acid
52.3 g	MOPS
33 g	Tris HCl
100 mL	Glucose (20% [w/v])

ZXY7 Media (q.s. 1000 mL)	
16 g	Tryptone
10 g	Yeast extract
5 g	Sodium chloride
25 mL	Glucose (20% [w/v])

SOC Media (q.s. 1000 mL)	
20 g	Tryptone
5 g	Yeast extract
700 mg	Sodium chloride
250 mg	Potassium chloride
10 mL	Magnesium sulfate (1 M)
18 mL	Glucose (20% [w/v])

Modified TB (q.s. 1000 mL)	
12 g	Tryptone
24 g	Yeast extract
5 g	Glycerol
5 g	Sodium chloride
25 mL	Glucose (20% [w/v])

Glucose M9Y Media (pH 6.9) (q.s. 1000 mL)	
6 g	Sodium phosphate dibasic
3 g	Potassium phosphate monobasic
5 g	Yeast extract
750 mg	Sodium chloride
1 mL	Calcium chloride (1 M)
1 mL	Magnesium sulfate (1 M)

Super Broth Media (pH 7.5) (q.s. 1000 mL)	
35 g	Tryptone
20 g	Yeast extract
1 g	Sodium chloride
25 mL	Glucose(20% [w/v])

TYGPN Media (pH 7.2) (q.s. 1000 mL)	
20 g	Tryptone
10 g	Yeast extract
6 g	Sodium phosphate dibasic
10 g	Potassium nitrate
25 mL	Glucose (20% [w/v])

NZCYM Media (pH 7.0) (q.s. 1000 mL)	
1 g	NZ-amine
14 g	Casamino acid
5 g	Yeast extract
8 g	Sodium chloride
800µL	Magnesium sulfate (1 M)
25 mL	Glucose (20% [w/v])

AI Media (q.s. 1000 mL)

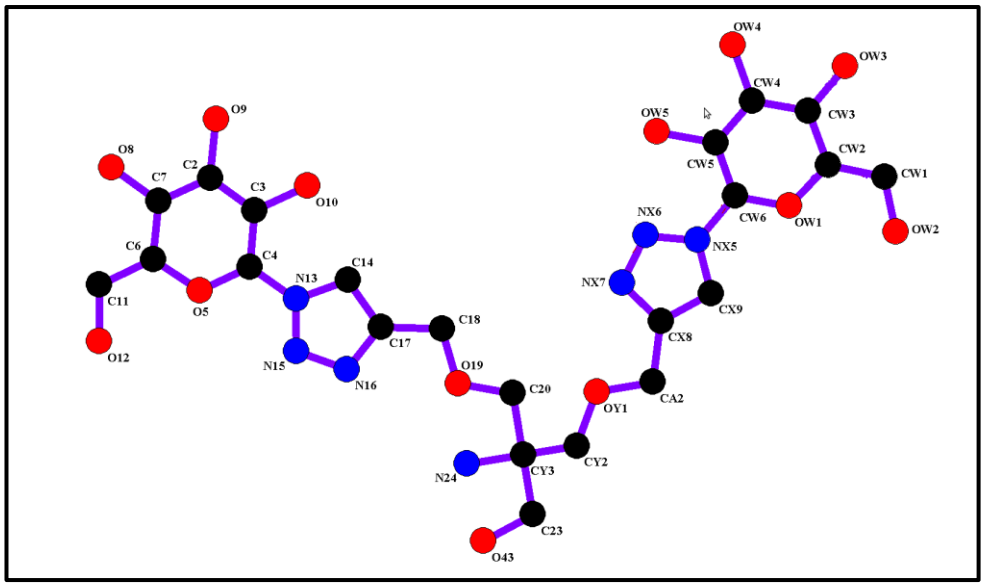
5 mL	Sodium sulfate (1 M)
2 mL	Magnesium sulfate (1 M)
200 µL	1000 x metals
20 mL	50x 5052
50 mL	20x NPS
900 mL	ZY media

50x 5052 (q.s. 100 mL)		**20x NPS (pH 6.8) (q.s. 100 mL)**	
25 g	Glycerol	5.4 g	Ammonium chloride
2.5 g	Glucose	6.8 g	Potassium phosphate (monobasic)
10 g	α-lactose	7.1 g	Sodium phosphate (dibasic)

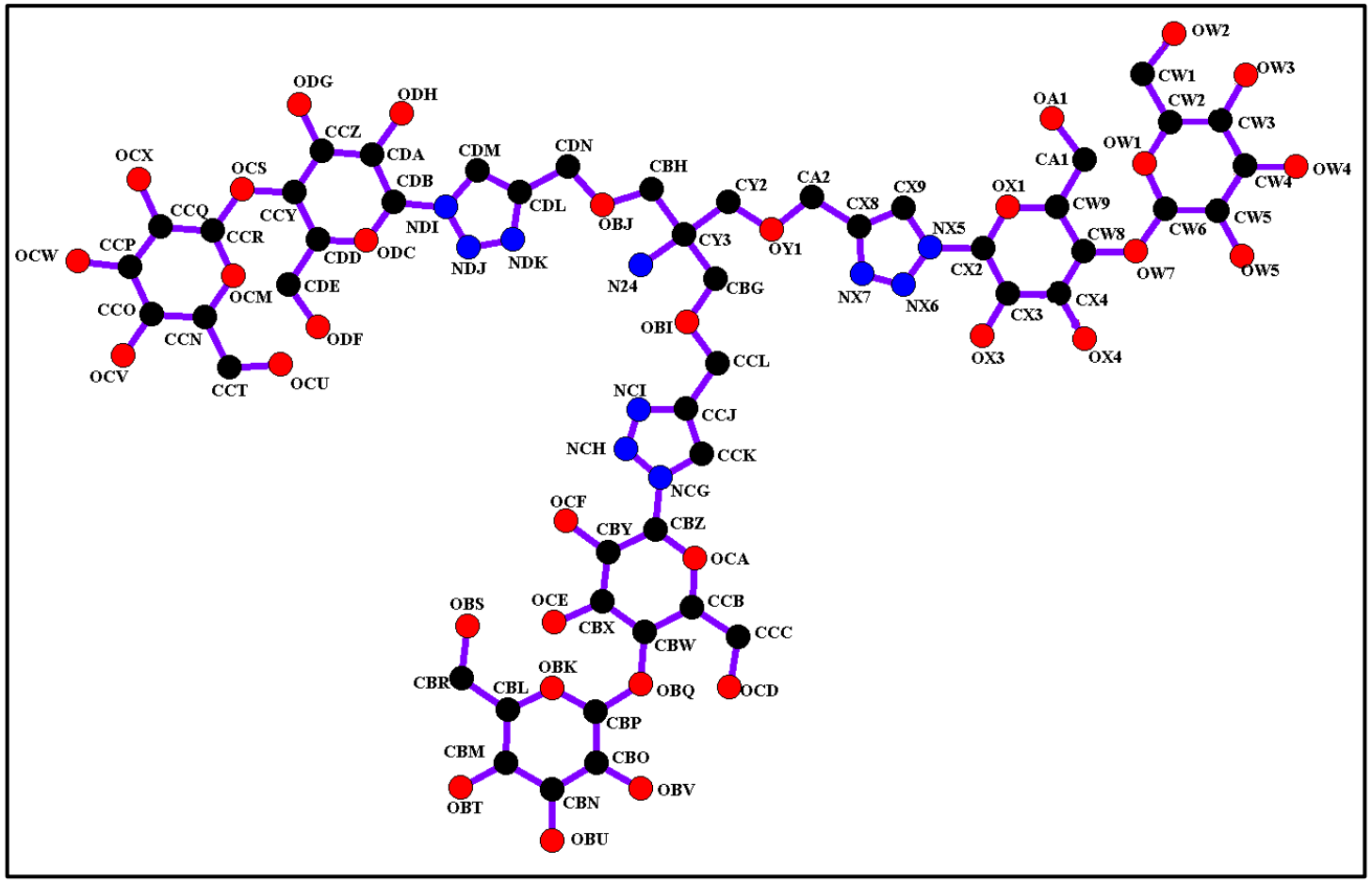
1000x metals (q.s. 100 mL)		**ZY media (q.s. 900 mL)**	
2 mL	Calcium chloride (1 M)	10 g	NZ-amine
1 mL	Manganese chloride dihydrate (1 M)	5 g	Yeast extract
1 mL	Zinc sulfate heptahydrate (1 M)		
1 mL	Cobalt chloride hexahydrate (0.2 M)		
2 mL	Copper chloride dehydrate (0.1 M)		
1 mL	Nickel chloride heptahydrate (0.2 M)		
2 mL	Sodium molybdate pentahydrate (0.1 M)		
2 mL	Sodium selenite pentahydrate (0.1 M)		
2 mL	Boric acid		
50 mL	Iron chloride hexahydrate (0.1 M in 0.1 M HCl)		

APPENDIX C: Atom labelling of dendrons

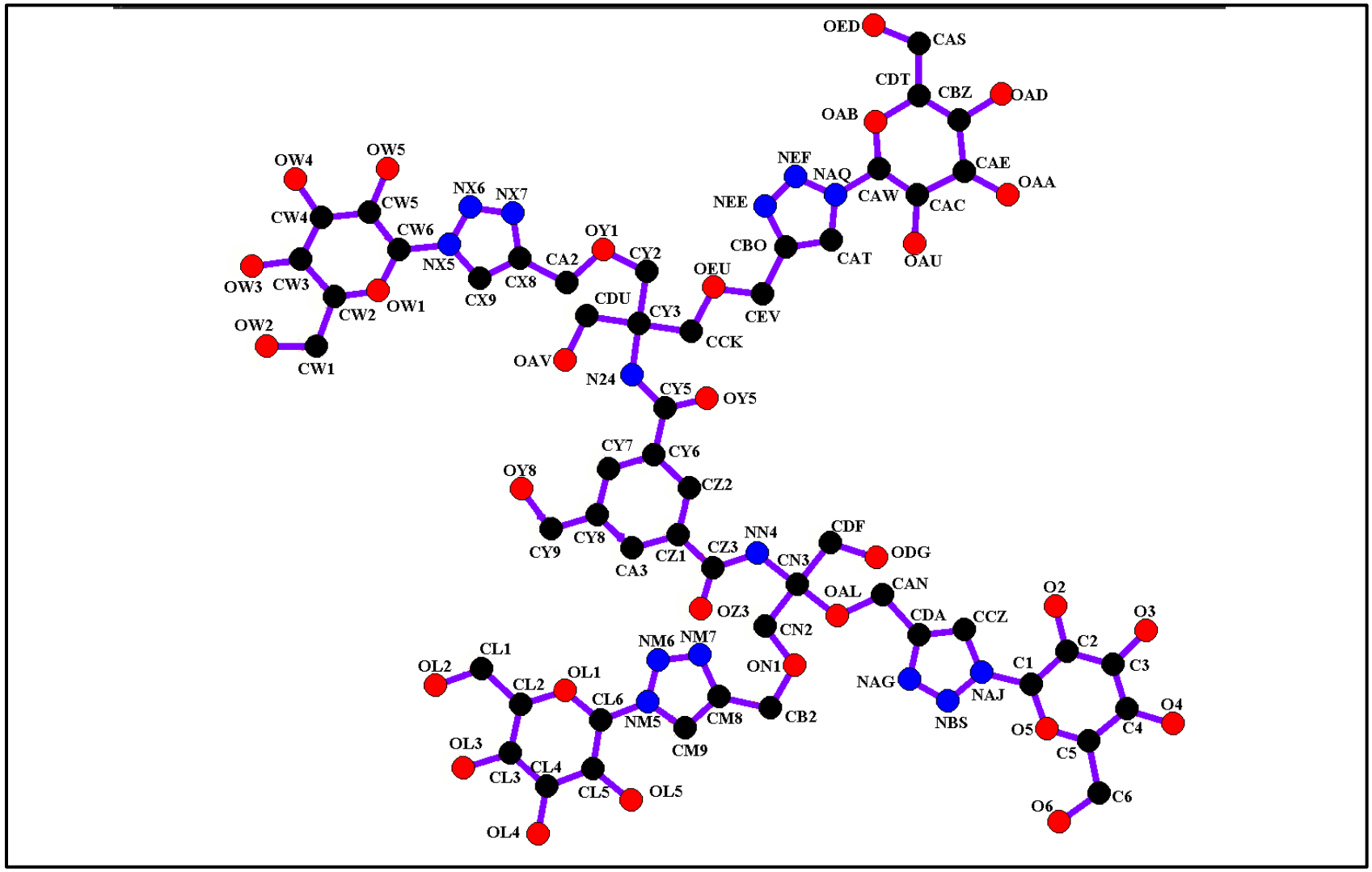
- **Dendron D1**



- **Dendron D2**

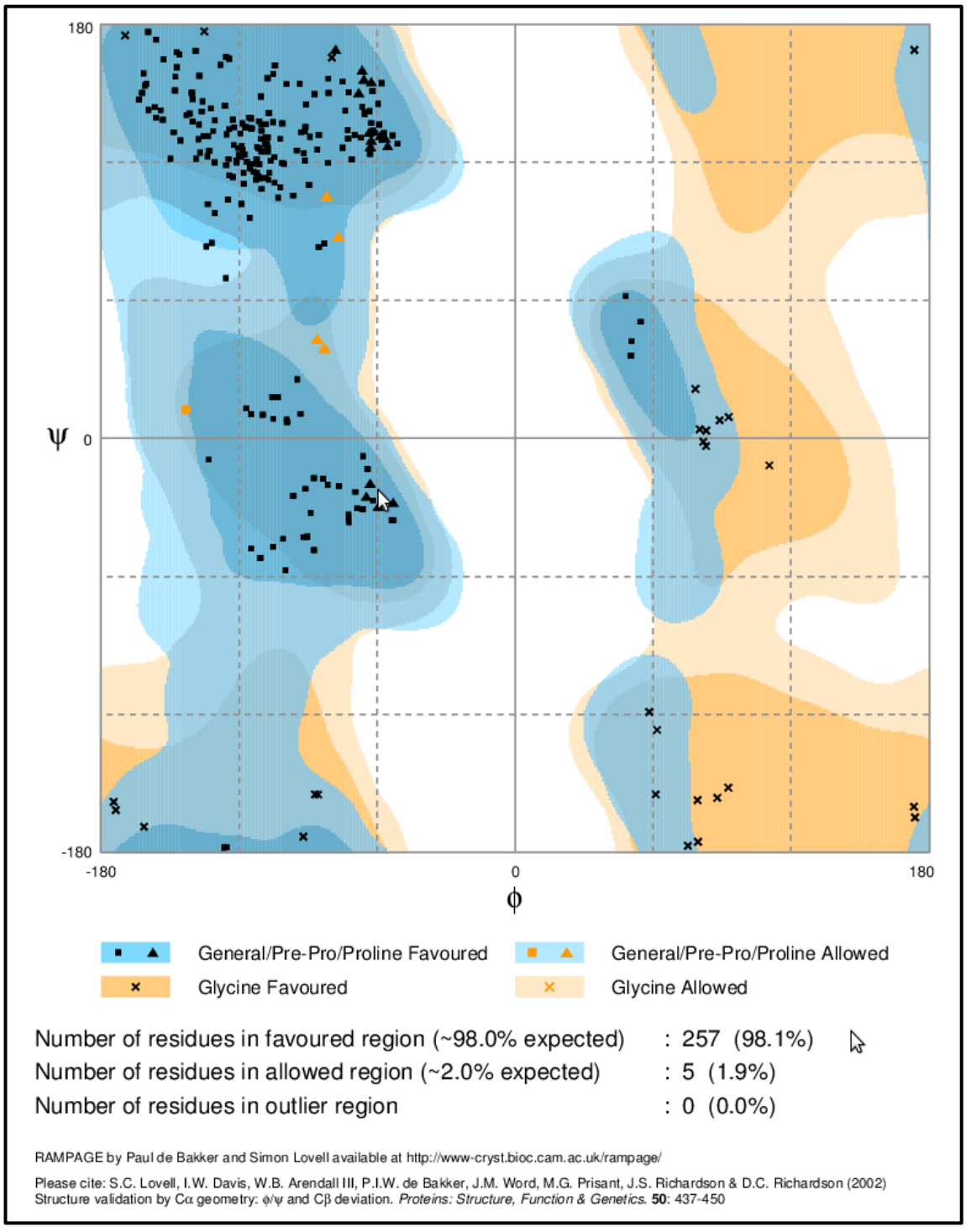

OW2
CW1
OW3
CW2
CW3
OA1
OW1
CA1
CW4
OW4
CW5
CW6
CW9
OX1
CW8
OW7
OW5
CX2
CX4
CX3
OX3
OX4
CX9
CX8
NX5
NX7
NX6
CA2
CY2
OY1
CY3
CBH
OBJ
CDN
CDL
CDM
NDI
NDJ
NDK
CDB
CDA
CDD
ODC
ODG
ODH
CCZ
CCY
OCS
OCX
CCQ
CCR
CCP
OCW
OCM
CDE
ODF
CCO
CCN
OCV
CCT
OCU
N24
CBG
OBI
CCL
NCI
CCJ
NCH
CCK
NCG
OCF
CBY
CBZ
OCA
OBS
OCE
CBX
CCB
CBW
CCC
OCD
CBR
CBL
OBK
OBQ
CBP
CBM
CBO
OBV
OBT
CBN
OBU

- **Dendron D3**

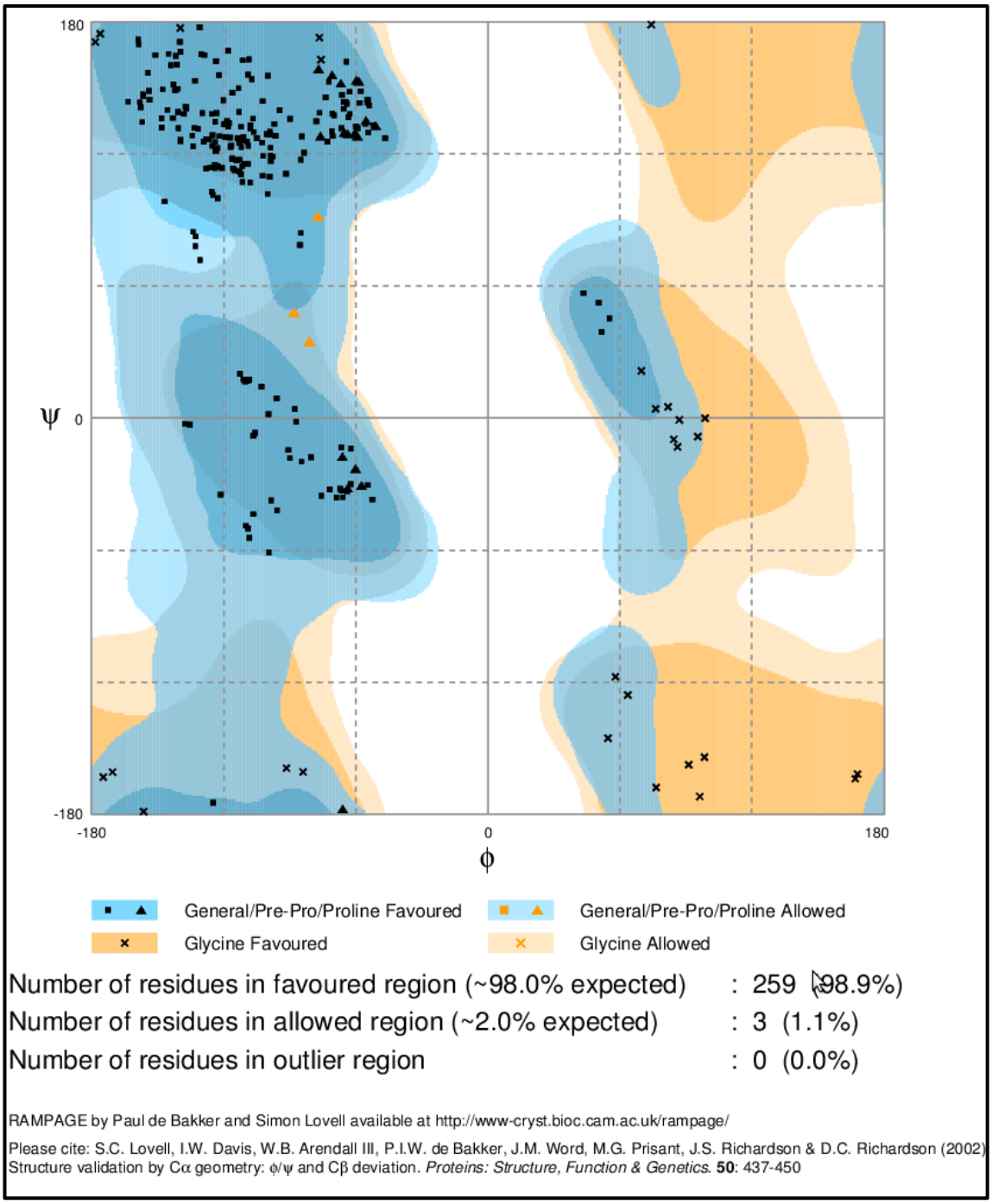

OED
CAS
CDT
CBZ
OAD
OAB
CAE
OAA
NEF
NAQ
CAW
CAC
NEE
OAU
CBO
CAT
OEU
CEV
CCK
CDU
CY3
CY2
OY1
CA2
CX8
CX9
NX5
NX6
NX7
CW6
CW5
CW4
OW4
OW5
OW3
CW3
CW2
OW1
OW2
CW1
OAV
N24
CY5
OY5
CY7
CY6
CZ2
OY8
CY8
CY9
CA3
CZ1
CZ3
NN4
CN3
CDF
ODG
OAL
CAN
OZ3
CN2
CDA
CCZ
O2
O3
C2
C3
C1
NAG
NAJ
NBS
O5
C4
O4
C5
C6
O6
NM6
NM7
ON1
CB2
CM8
NM5
CM9
CL6
CL1
OL2
CL2
OL1
OL3
CL3
CL4
CL5
OL5
OL4

APPENDIX D: Ramachandran plots

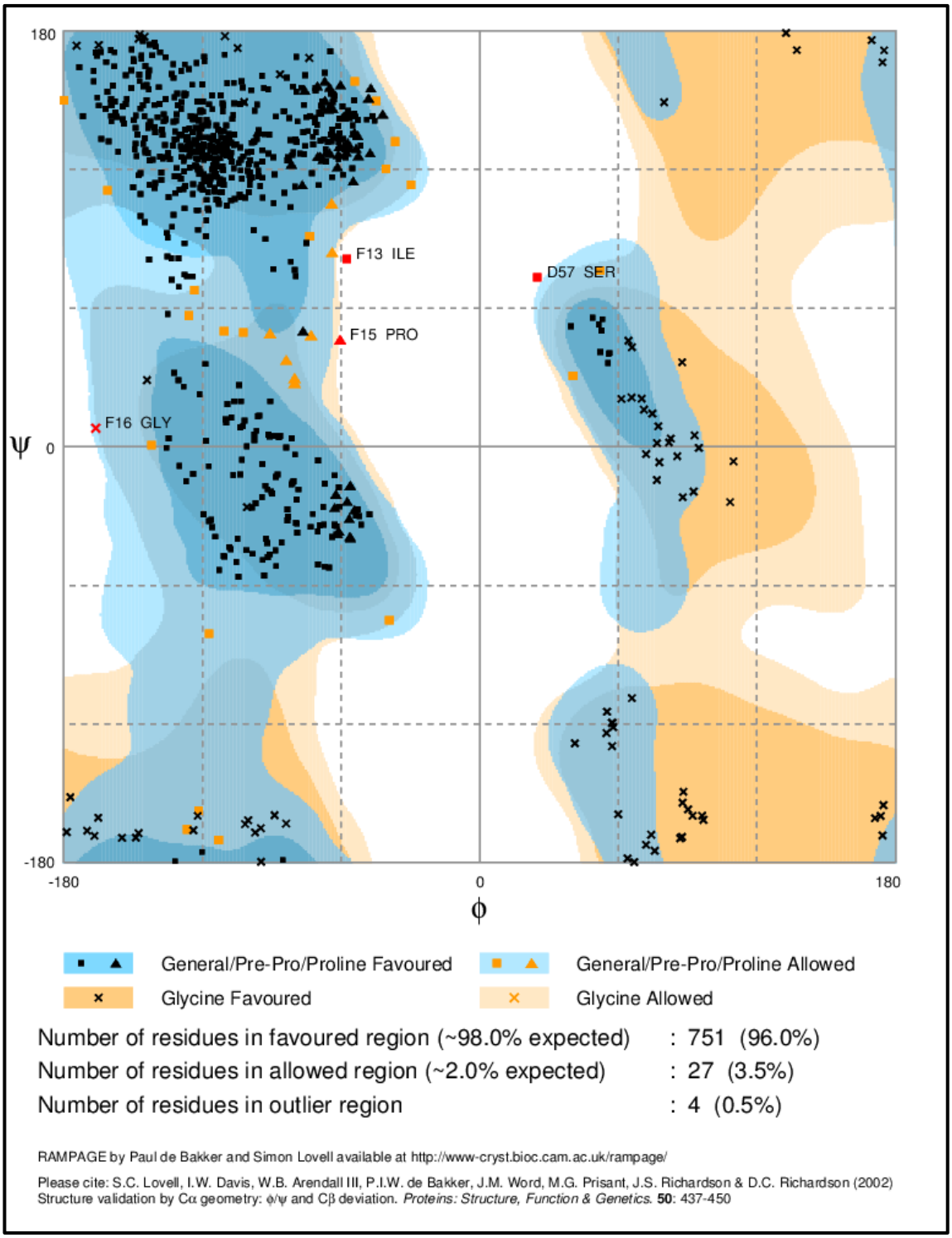
- **hGal-7 in complex with Dendron D1**



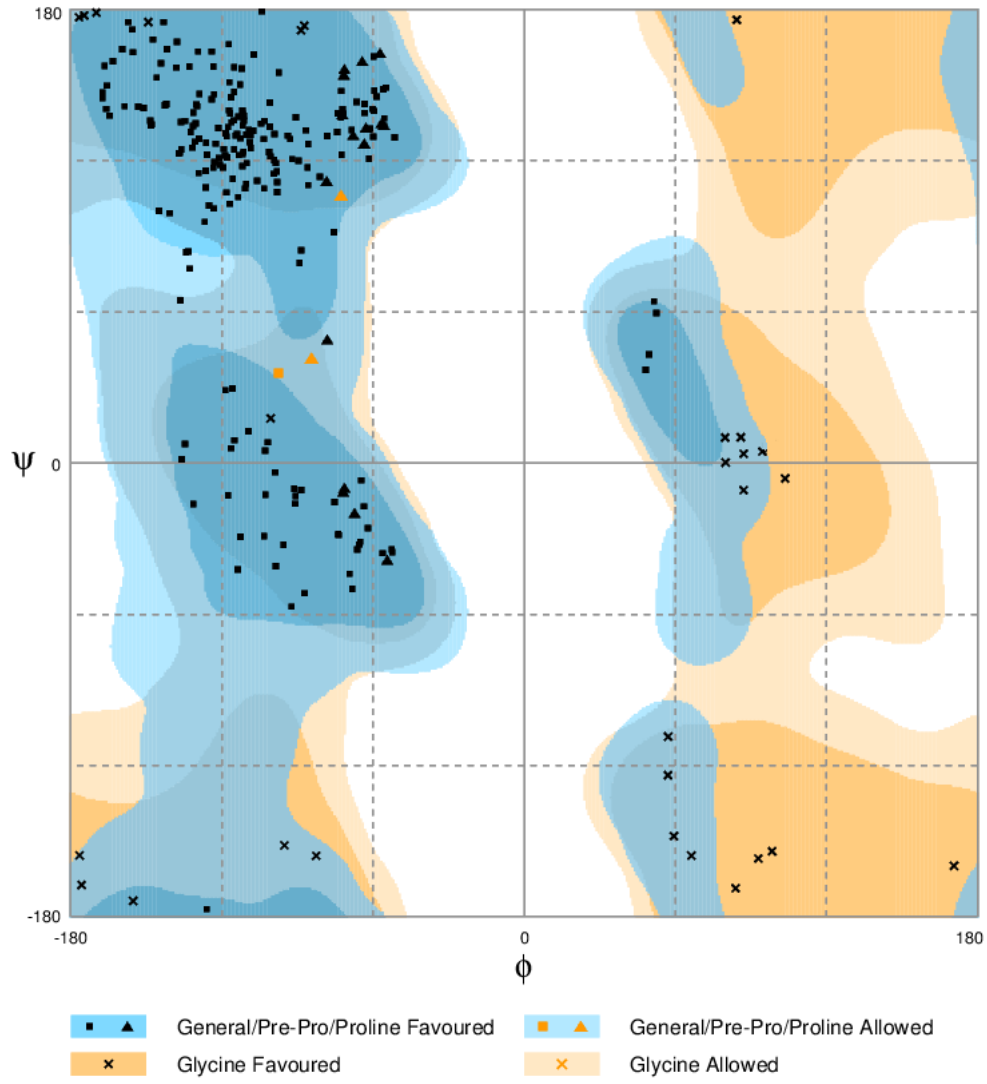
- **hGal-7 in complex with Dendron D2-1**



- **hGal-7 in complex with Dendron D2-2**



- **hGal-7 in complex with Dendron D3**



Number of residues in favoured region (~98.0% expected) : 256 (98.8%)
Number of residues in allowed region (~2.0% expected) : 3 (1.2%)
Number of residues in outlier region : 0 (0.0%)

RAMPAGE by Paul de Bakker and Simon Lovell available at http://www-cryst.bioc.cam.ac.uk/rampage/

Please cite: S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson & D.C. Richardson (2002) Structure validation by Cα geometry: ϕ/ψ and Cβ deviation. *Proteins: Structure, Function & Genetics*. **50**: 437-450

APPENDIX E: Protein sequences

Appendix E1: SIRT2

E1.1: Protein sequence alignment for SIRT2.6a. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT2.6a with the expected SIRT2 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          MAHHHHHHVDDDDKEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVI 60
Expected        MAHHHHHHVDDDDKEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVI 60
                ************************************************************

Result          CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 120
Expected        CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 120
                ************************************************************

Result          GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 180
Expected        GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 180
                ************************************************************

Result          RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 240
Expected        RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 240
                ************************************************************

Result          GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKK------- 293
Expected        GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAW 300
                *************************************************

Result          ---------------------------------------------------------
Expected        LGECDQGCLALAELLGWKKELEDLVRREHASIDAQSGAGVPNPSTSASPKKSPPPAKDEA 360


Result          -------
Expected        RTKEKPQ 367
```

E1.2: Protein sequence alignment for SIRT2.6a. Alignment of the translated DNA sequencing result after reverse sequencing (T7 term primer) of the expression vector carrying the gene of interest SIRT2.6a with the expected SIRT2 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        MAHHHHHHVDDDDKEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVI 60


Result          CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 60
Expected        CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 120
                ************************************************************


Result          GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 120
Expected        GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 180
                ************************************************************


Result          RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 180
Expected        RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 240
                ************************************************************


Result          GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAW 240
Expected        GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAW 300
                ************************************************************


Result          LGECDQGCLALAELLGWKKELEDLVRREHASIDAQSGAGVPNPSTSASPKKSPPPAKDEA 300
Expected        LGECDQGCLALAELLGWKKELEDLVRREHASIDAQSGAGVPNPSTSASPKKSPPPAKDEA 360
                ************************************************************


Result          RTKEKPQ-TGLLLKSRVW-RNRCCEI-TPAHGLVY-RS- 334
Expected        RTKEKPQ------------------------------- 367
```

E1.3: Protein sequence alignment for SIRT2.1. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT2.1 with the expected SIRT2 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ELSGTFNSTQHNIL-LNKNYYQIICILIKILYCKLHFIYNQRRYTMAHHHHHHSSGKETA 59
Expected        ----------------------------------------------MAHHHHHHSSGKETA 15
                                                              ***************


Result          AAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMEADMDFLRNLFSQTLSLGSQKERLLD 119
Expected        AAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMEADMDFLRNLFSQTLSLGSQKERLLD 75
                ************************************************************


Result          ELTLEGVARYMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFE 179
Expected        ELTLEGVARYMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFE 135
                ************************************************************


Result          ISYFKKHPEPFFALAKELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQ 239
Expected        ISYFKKHPEPFFALAKELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQ 195
                ************************************************************


Result          EDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKIFSE------------------------ 296
Expected        EDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLP 255
                ********************************


Result          ------------------------------------------------------------
Expected        ARFFSCMQSDFLKVDLLLVMGTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMI 315


Result          ------------------------------------------------------
Expected        MGLGGGMDFDSKKAYRDVAWLGECDQGCLALAELLGWKKELEDLVRREHASIDAQS 371
```

E1.4: Protein sequence alignment for SIRT2.1. Alignment of the translated DNA sequencing result after reverse sequencing (TriExDOWN primer) of the expression vector carrying the gene of interest SIRT2.1 with the expected SIRT2 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMEADMDFLRNLF 60


Result          ------------------------------------------------------------
Expected        SQTLSLGSQKERLLDELTLEGVARYMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYD 120


Result          ------------------------------------------------------------
Expected        NLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYPGQFKPTICHYFMRLLKDKGLLLRCY 180


Result          ------------------------------SASCRHEYPLSWMKEKIFSEVTPKCEDCQ 29
Expected        TQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKIFSEVTPKCEDCQ 240
                                              ****************************


Result          SLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVMGTSLQVQPFASLISKAPLSTPRLLI 89
Expected        SLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVMGTSLQVQPFASLISKAPLSTPRLLI 300
                ************************************************************


Result          NKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAELLGWKKELEDLV 149
Expected        NKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAELLGWKKELEDLV 360
                ************************************************************


Result          RREHASIDAQS-PGFSSTISELVDPNSQISARLQVDGTGSKLAAAQLYTRASQPELAPED 208
Expected        RREHASIDAQS---------------------------------------------------- 371
                ***********


Result          PEDLEHHHHHHHH-VINLRCRLPIRRWW 235
Expected        ----------------------------
```

E1.5: Protein sequence alignment for SIRT2.4. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT2.4 with the expected SIRT2 protein sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          LE-FCLL-EGDIHMEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVI 58
Expected        ------------MEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVI 47
                            ********************************************


Result          CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 118
Expected        CLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYP 107
                ************************************************************


Result          GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 178
Expected        GQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASC 167
                ************************************************************


Result          RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 238
Expected        RHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVM 227
                ************************************************************


Result          GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKPTG--TW 296
Expected        GTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAW 287
                ************************************************.    :*


Result          P------------------------------------------- 297
Expected        LGECDQGCLALAELLGWKKELEDLVRREHASIDAQSLEHHHHHH 331
```

E1.6: Protein sequence alignment for SIRT2.4. Alignment of the translated DNA sequencing result after forward sequencing (T7 term primer) of the expression vector carrying the gene of interest SIRT2.4 with the expected SIRT2 protein sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ----------ASAGRADLG-RGGPVHAERTLS------QSH------LLVGAGISTSAGI 37
Expected        MEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERCRRVICLVGAGISTSAGI 60
                           :   .** :   :  * **.      **.       ************

Result          PDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYPGQFKPTICHYFMR 97
Expected        PDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYPGQFKPTICHYFMR 120
                ************************************************************

Result          LLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKI 157
Expected        LLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKI 180
                ************************************************************

Result          FSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVMGTSLQVQPFASLI 217
Expected        FSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLLVMGTSLQVQPFASLI 240
                ************************************************************

Result          SKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAE 277
Expected        SKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAE 300
                ************************************************************

Result          LLGWKKELEDLVRREHASIDAQSLEHHHHHH-DPAANKAR 316
Expected        LLGWKKELEDLVRREHASIDAQSLEHHHHHH--------- 331
                ******************************
```

E1.7: Protein sequence alignment for SIRT2.7. Alignment of the translated DNA sequencing result after forward sequencing [M13(-21) primer] of the PCR blunt vector carrying the gene of interest SIRT2.7 with the expected SIRT2 protein sequence using clustal 2.1.

```
Result          -GDWAL-MHARAAASVMDICRIQGSEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVAR 58
Expected        ------------------------EADMDFLRNLFSQTLSLGSQKERLLDELTLEGVAR 35
                                        **********************************

Result          YMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPE 118
Expected        YMQSERCRRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPE 95
                ************************************************************

Result          PFFALAKELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGT 178
Expected        PFFALAKELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGT 155
                ************************************************************

Result          FYTSHCVSASCRHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQS 238
Expected        FYTSHCVSASCRHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQS 215
                ************************************************************

Result          DFLKVDLLLVMGTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLG----- 293
Expected        DFLKVDLLLVMGTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDF 275
                ************************************************

Result          ---------------------------------------------
Expected        DSKKAYRDVAWLGECDQGCLALAELLGWKKELEDLVRREHASIDAQS 322
```

E1.8: Protein sequence alignment for SIRT2.7. Alignment of the translated DNA sequencing result after forward sequencing (pGEx forward primer) of the expression vector carrying the gene of interest SIRT2.7 with the expected SIRT2 protein sequence using clustal 2.1.

```
Result          ATSSKSDLEVLFQGPLGSEADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERC 60
Expected        ------------------EADMDFLRNLFSQTLSLGSQKERLLDELTLEGVARYMQSERC 42
                                  ****************************************

Result          RRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAK 120
Expected        RRVICLVGAGISTSAGIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAK 102
                ************************************************************

Result          ELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCV 180
Expected        ELYPGQFKPTICHYFMRLLKDKGLLLRCYTQNIDTLERIAGLEQEDLVEAHGTFYTSHCV 162
                ************************************************************

Result          SASCRHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDL 240
Expected        SASCRHEYPLSWMKEKIFSEVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDL 222
                ************************************************************

Result          LLVMGTSLQVQPFASLISKAPLSTPRLLINKGES-------------------------- 289
Expected        LLVMGTSLQVQPFASLISKAPLSTPRLLINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYR 282
                ****************************** ::

Result          ------------------------------------ 299
Expected        DVAWLGECDQGCLALAELLGWKKELEDLVRREHASIDAQS 322
```

Appendix E2: SIRT7

E2.1: Protein sequence alignment for SIRT7.1. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT7.1 with the expected SIRT7 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          RS-SGTFNSTQHNIL-LNKNYYQIICILIKILYCKLHFIYNQRRYTMAHHHHHHSSGKET 58
Expected        ----------------------------------------------MAHHHHHHSSGKET 14
                                                              **************

Result          AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMAAGGLSRSERKAAERVRRLREEQQR 118
Expected        AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMAAGGLSRSERKAAERVRRLREEQQR 74
                ************************************************************

Result          ERLRQVSRILRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELR 178
Expected        ERLRQVSRILRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELR 134
                ************************************************************

Result          GKVRELASAVRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPT 238
Expected        GKVRELASAVRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPT 194
                ************************************************************

Result          LTHMSITRLHEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVR 298
Expected        LTHMSITRLHEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVR 254
                ************************************************************

Result          VFDVTERTALHRHQTG---------------------------------------------- 314
Expected        VFDVTERTALHRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTILC 314
                ****************

Result          ------------------------------------------------------------
Expected        LGSSLKVLKKYPRLWCMTKPPSRRPKLYIVNLQWTPKDDWAALKLHGKCDDVMRLLMAEL 374


Result          ------------------------------------------------------------
Expected        GLEIPAYSRWQDPIFSLATPLRAGEEGSHSRKSLCRSREEAPPGDRGAPLSSAPILGGWF 434


Result          --------------
Expected        GRGCTKRTKRKKVT 448
```

E2.2: Protein sequence alignment for SIRT7.6. Alignment of the translated DNA sequencing result after forward sequencing (S-tag primer) of the expression vector carrying the gene of interest SIRT7.6 with the expected SIRT7 sequence using clustal 2.1. GST-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDG 60

Result          ------------------------------------------------------------
Expected        DVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVD 120

Result          ------------------------------------------------------------
Expected        FLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKK 180

Result          ------------------------------------------------------------
Expected        RIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDGSTSGSGHHHHHHSAGLVPRG 240

Result          ---------------------------GGSGDDDDKMAAGGLSRSERKAAERVRRLREEQ 33
Expected        STAIGMKETAAAKFERQHMDSPDLGTGGGSGDDDDKMAAGGLSRSERKAAERVRRLREEQ 300
                                           ********************************

Result          QRERLRQVSRILRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEE 93
Expected        QRERLRQVSRILRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEE 360
                ************************************************************

Result          LRGKVRELASAVRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAE 153
Expected        LRGKVRELASAVRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAE 420
                ************************************************************

Result          PTLTHMSITRLHEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREY 213
Expected        PTLTHMSITRLHEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREY 480
                ************************************************************

Result          VRVFDVTERTALHRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTI 273
Expected        VRVFDVTERTALHRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTI 540
                ************************************************************

Result          LCLGSSLKGSKEVPTPLVHDQAP----------------------------------- 296
Expected        LCLGSSLKVLKKYPRLWCMTKPPSRRPKLYIVNLQWTPKDDWAALKLHGKCDDVMRLLMA 600
                ********  *: *      :.*

Result          ------PAA---------------------------------------------------- 299
Expected        ELGLEIPAYSRWQDPIFSLATPLRAGEEGSHSRKSLCRSREEAPPGDRGAPLSSAPILGG 660
                      **

Result          ----------------
Expected        WFGRGCTKRTKRKKVT 676
```

E2.3: Protein sequence alignment for SIRT7.7. Alignment of the translated DNA sequencing result after forward sequencing [M13(-21) primer] of the PCR blunt vector carrying the gene of interest SIRT7.7 with the expected SIRT7 sequence using clustal 2.1.

```
Result          RAIGPSRCMLERPPV-WISAEFREFMAAGGLSRSERKAAERVRRLREEQQRERLRQVSRI 59
Expected        -------------------------MAAGGLSRSERKAAERVRRLREEQQRERLRQVSRI 35
                                         **********************************

Result          LRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELRGKVRELASA 119
Expected        LRKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELRGKVRELASA 95
                ************************************************************

Result          VRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPTLTHMSITRL 179
Expected        VRNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPTLTHMSITRL 155
                ************************************************************

Result          HEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVRVFDVTERTA 239
Expected        HEQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVRVFDVTERTA 215
                ************************************************************

Result          LHRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLN------------------------- 274
Expected        LHRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTILCLGSSLKVLK 275
                **********************************

Result          ------------------------------------------------------------
Expected        KYPRLWCMTKPPSRRPKLYIVNLQWTPKDDWAALKLHGKCDDVMRLLMAELGLEIPAYSR 335


Result          ------------------------------------------------------------
Expected        WQDPIFSLATPLRAGEEGSHSRKSLCRSREEAPPGDRGAPLSSAPILGGWFGRGCTKRTK 395


Result          -----
Expected        RKKVT 400
```

E2.4: Protein sequence alignment for SIRT7.7. Alignment of the translated DNA sequencing result after forward sequencing (pGEx forward primer) of the expression vector carrying the gene of interest SIRT7.7 with the expected SIRT7 sequence using clustal 2.1.

```
Result          DHPPKSDLEVLFQGPLGSPEFREFMAAGGLSRSERKAAERVRRLREEQQRERLRQVSRIL 60
Expected        ------------------------MAAGGLSRSERKAAERVRRLREEQQRERLRQVSRIL 36
                                         **********************************

Result          RKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELRGKVRELASAV 120
Expected        RKAAAERSAEEGRLLAESADLVTELQGRSRRREGLKRRQEEVCDDPEELRGKVRELASAV 96
                ************************************************************

Result          RNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPTLTHMSITRLH 180
Expected        RNAKYLVVYTGAGISTAASIPDYRGPNGVWTLLQKGRSVSAADLSEAEPTLTHMSITRLH 156
                ************************************************************

Result          EQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVRVFDVTERTAL 240
Expected        EQKLVQHVVSQNCDGLHLRSGLPRTAISELHGNMYIEVCTSCVPNREYVRVFDVTERTAL 216
                ************************************************************

Result          HRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTILCLGSSLKVLKK 300
Expected        HRHQTGRTCHKCGTQLRDTIVHFGERGTLGQPLNWEAATEAASRADTILCLGSSLKVLKK 276
                ************************************************************

Result          YPRLWCMTKPPSRRPKLYIVNLQWTPKDDWAALKLHGKCDDVMR---------------- 344
Expected        YPRLWCMTKPPSRRPKLYIVNLQWTPKDDWAALKLHGKCDDVMRLLMAELGLEIPAYSRW 336
                ****************************************

Result          ------------------------------------------------------------
Expected        QDPIFSLATPLRAGEEGSHSRKSLCRSREEAPPGDRGAPLSSAPILGGWFGRGCTKRTKR 396


Result          ----
Expected        KKVT 400
```

Appendix E3: SIRT1

E3.1: Protein sequence alignment for SIRT1.1. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.1 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          PELSGTFNSTQHNIL-LNKNYYQIICILIKILYCKLHFIYNQRRYTMAHHHHHHSSGKET 59
Expected        ----------------------------------------------MAHHHHHHSSGKET 14
                                                              **************

Result          AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMIGYRDNLLFGDEIITNGFHSCESDE 119
Expected        AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMIGYRDNLLFGDEIITNGFHSCESDE  74
                ************************************************************

Result          EDRASHASSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIV 179
Expected        EDRASHASSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIV 134
                ************************************************************

Result          INILSEPPKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLA 239
Expected        INILSEPPKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLA 194
                ************************************************************

Result          VDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPAQFQPSLCHKFIALSDKG----RKTT 295
Expected        VDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYT 254
                ********************************.***************     *: *

Result          SQLYPEHRHAGTG----CGN---------------------------------PKDNSV 317
Expected        QNIDTLEQVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPL 314
                .:: . .: **      * .                                 * *:.:

Result          SWF------------------------------------------------------- 320
Expected        AIMKPEIVFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILI 374
                : :

Result          ---------------------------------------------------------
Expected        NREPLPHLHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYL 434


Result          ---------------------------------------------------------
Expected        SELPPTPLHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQT 494


Result          ---------------------------------------------------------
Expected        SRNVESIAEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLF 554


Result          ---------------------------------------------------------
Expected        LPPNRYIFHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLED 614


Result          -------------------------------------------
Expected        EPDVPERAGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS 657
```

E3.2: Protein sequence alignment for SIRT1.1. Alignment of the translated DNA sequencing result after reverse sequencing (TriExDOWN primer) of the expression vector carrying the gene of interest SIRT1.1 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMIGYRDNLLFGD 60


Result          ------------------------------------------------------------
Expected        EIITNGFHSCESDEEDRASHASSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETI 120


Result          ------------------------------------------------------------
Expected        PPPELDDMTLWQIVINILSEPPKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGI 180


Result          ------------------------------------------------------------
Expected        PDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFI 240


Result          ------------------------------------------------------------
Expected        ALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQ 300


Result           -------------------------------------------LLHVNGSSS-SKTSSI 15
Expected        VVPRCPRCPADEPLAIMKPEIVFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVAL 360
                                                           ** * ***   :. ::

Result          IPSPYPMKCLRY--LIENLASSAFDVELLRL--CHINELLS-VRW-ICQTCCNPVKLSEI 69
Expected        IPSSIPHEVPQILINREPLPHLHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEI 420
                ***. * :  :     * *.   ******      ****   :    .: **********

Result          TEKPPRTQKELAYLSELPPTPLHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSE 129
Expected        TEKPPRTQKELAYLSELPPTPLHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSE 480
                ************************************************************

Result          SKGCMEEKPQEVQTSRNVESIAEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAK 189
Expected        SKGCMEEKPQEVQTSRNVESIAEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAK 540
                ************************************************************

Result          EQISRRLDGNQYLFLPPNRYIFHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPME 249
Expected        EQISRRLDGNQYLFLPPNRYIFHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPME 600
                ************************************************************

Result          DESEIEEFYNGLEDEPDVPERAGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS-TG 308
Expected        DESEIEEFYNGLEDEPDVPERAGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS--- 657
                ********************************************************

Result          LLLNDI-ARGSEFSDLGAPAGRRYRFEACGRTAVYTCKPARTRPGRPRGSRAPPSPSPSL 367
Expected        ------------------------------------------------------------


Result          SD-PQVQAAYQKVVAVAQVL 386
Expected        --------------------
```

E3.3: Protein sequence alignment for SIRT1.5. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.5 with the expected SIRT1 sequence using clustal 2.1. GST-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDG 60


Result          ------------------------------------------------------------
Expected        DVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVD 120


Result          ------------------------------------------------------------
Expected        FLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKK 180


Result          ----------------WWFW---------------------------------------- 4
Expected        RIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDGSTSGSGHHHHHHSAGLVPRG 240
                                :  *

Result          --------------------------------RRQMIGYRDNLLFGDEIITNGFHSCESD 32
Expected        STAIGMKETAAAKFERQHMDSPDLGTGGGSGDDDDKIGYRDNLLFGDEIITNGFHSCESD 300
                                               : ************************

Result          EEDRASHASSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQI 92
Expected        EEDRASHASSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQI 360
                ************************************************************

Result          VINILSEPPKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARL 152
Expected        VINILSEPPKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARL 420
                ************************************************************

Result          AVDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNY 212
Expected        AVDFPDLPDPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNY 480
                ************************************************************

Result          TQNIDTLEQVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEP 272
Expected        TQNIDTLEQVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEP 540
                ************************************************************

Result          LAIMKPEIVFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALISKF--------- 323
Expected        LAIMKPEIVFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQIL 600
                ********************************************..

Result          ------------------------------------------------------------
Expected        INREPLPHLHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAY 660


Result          ------------------------------------------------------------
Expected        LSELPPTPLHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQ 720


Result          ------------------------------------------------------------
Expected        TSRNVESIAEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYL 780


Result          ------------------------------------------------------------
Expected        FLPPNRYIFHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLE 840


Result          ---------------------------------------------
Expected        DEPDVPERAGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS- 884
```

E3.4: Protein sequence alignment for SIRT1.6. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.6 with the expected SIRT1 sequence using clustal 2.1. GST-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDG 60


Result          ------------------------------------------------------------
Expected        DVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVD 120


Result          ------------------------------------------------------------
Expected        FLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKK 180


Result          ------------------------------------------------------------
Expected        RIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDGSTSGSGHHHHHHSAGLVPRG 240


Result          ---------------------------GGSGDDDDKMCESDEEDRASHASSSDWTPRPRI 33
Expected        STAIGMKETAAAKFERQHMDSPDLGTGGGSGDDDDKMCESDEEDRASHASSSDWTPRPRI 300
                                           ********************************


Result          GPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINT 93
Expected        GPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINT 360
                ************************************************************


Result          IEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYF 153
Expected        IEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYF 420
                ************************************************************


Result          RKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQC 213
Expected        RKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQC 480
                ************************************************************


Result          HGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQF 273
Expected        HGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQF 540
                ************************************************************


Result          HRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLG-- 331
Expected        HRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLGDC 600
                ************************************************************


Result          ------------------------------------------------------------
Expected        DVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTPLHVSEDSSSPE 660


Result          ------------------------------------------------------------
Expected        RTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESIAEQMENPDLKN 720


Result          ------------------------------------------------------------
Expected        VGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYIFHGAEVYSDSE 780


Result          ------------------------------------------------------------
Expected        DDVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLEDEPDVPERAGGAGFGTDGD 840


Result          -------------------------
Expected        DQEAINEAISVKQEVTDMNYPSNKS 865
```

E3.5: Protein sequence alignment for SIRT1.7. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.7 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          FPELSGTFNSTQHNIL-LNKNYYQIICILIKILYCKLHFIYNQRRYTMAHHHHHHSSGKE 59
Expected        ----------------------------------------------MAHHHHHHSSGKE 13
                                                             *************

Result          TAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMCESDEEDRASHASSSDWTPRPRIG 119
Expected        TAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMCESDEEDRASHASSSDWTPRPRIG 73
                ************************************************************

Result          PYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINTI 179
Expected        PYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINTI 133
                ************************************************************

Result          EDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFR 239
Expected        EDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFR 193
                ************************************************************

Result          KDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQCH 299
Expected        KDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQCH 253
                ************************************************************

Result          G------------------------------------------------------------ 300
Expected        GSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQFH 313
                *

Result          ------------------------------------------------------------
Expected        RAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLGDCD 373


Result          ------------------------------------------------------------
Expected        VIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTPLHVSEDSSSPER 433


Result          ------------------------------------------------------------
Expected        TSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESIAEQMENPDLKNV 493


Result          ------------------------------------------------------------
Expected        GSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYIFHGAEVYSDSED 553


Result          ------------------------------------------------------------
Expected        DVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLEDEPDVPERAGGAGFGTDGDD 613


Result          ------------------------
Expected        QEAINEAISVKQEVTDMNYPSNKS 637
```

E3.6: Protein sequence alignment for SIRT1.7. Alignment of the translated DNA sequencing result after reverse sequencing (TriExDOWN primer) of the expression vector carrying the gene of interest SIRT1.7 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKCESDEEDRASHA 60


Result          ------------------------------------------------------------
Expected        SSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEP 120


Result          ------------------------------------------------------------
Expected        PKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLP 180


Result          ------------------------------------------------------------
Expected        DPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLE 240


Result          ------------------------------------------------------------
Expected        QVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEI 300


Result          ------------------------------------------------------------
Expected        VFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPH 360


Result          ------------------------EYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTP 36
Expected        LHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTP 420
                                         **********************************

Result          LHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESI 96
Expected        LHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESI 480
                ************************************************************

Result          AEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYI 156
Expected        AEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYI 540
                ************************************************************

Result          FHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLEDEPDVPER 216
Expected        FHGAEVYSDSEDDVLSSSSCGSNSDSGTCQSPSLEEPMEDESEIEEFYNGLEDEPDVPER 600
                ************************************************************

Result          AGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS-PGFSSTISELVDPNSQISARLQV 275
Expected        AGGAGFGTDGDDQEAINEAISVKQEVTDMNYPSNKS------------------------ 636
                **********************************

Result          DGTGSKLAAAQLYTRASQPELAPEDPEDLEHHHHHHHH-VINLRCRLPIRRWW 327
Expected        ------------------------------------------------------
```

E3.7: Protein sequence alignment for SIRT1.8. Alignment of the translated DNA sequencing result after reverse sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.8 with the expected SIRT1 sequence using clustal 2.1. GST-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        SPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDG 60


Result          ------------------------------------------------------------
Expected        DVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVD 120


Result          ------------------------------------------------------------
Expected        FLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKK 180


Result          ------------------------------------------------------------
Expected        RIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDGSTSGSGHHHHHHSAGLVPRG 240


Result          ---------------------------GGSGDDDDKMCESDEEDRASHASSSDWTPRPRI 33
Expected        STAIGMKETAAAKFERQHMDSPDLGTGGGSGDDDDKMCESDEEDRASHASSSDWTPRPRI 300
                                           ********************************

Result          GPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINT 93
Expected        GPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINT 360
                ************************************************************

Result          IEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYF 153
Expected        IEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYF 420
                ************************************************************

Result          RKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQC 213
Expected        RKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQC 480
                ************************************************************

Result          HGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQF 273
Expected        HGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQF 540
                ************************************************************

Result          HRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLG-- 331
Expected        HRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLGDC 600
                ************************************************************

Result          ------------------------------------------------------------
Expected        DVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTPLHVSEDSSSPE 660


Result          ------------------------------------------------------------
Expected        RTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESIAEQMENPDLKN 720


Result          ------------------------------------------------------------
Expected        VGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYIFHGAEVYSDSE 780


Result          ---
Expected        DDV 783
```

E3.8: Protein sequence alignment for SIRT1.9. Alignment of the translated DNA sequencing result after forward sequencing (T7 primer) of the expression vector carrying the gene of interest SIRT1.9 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          PELSGTFNSTQHNIL-LNKNYYQIICILIKILYCKLHFIYNQRRYTMAHHHHHHSSGKET 59
Expected        ---------------------------------------------MAHHHHHHSSGKET 14
                                                             **************

Result          AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMCESDEEDRASHASSSDWTPRPRIGP 119
Expected        AAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKMCESDEEDRASHASSSDWTPRPRIGP 73
                ************************************************************

Result          YTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINTIE 179
Expected        YTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEPPKRKKRKDINTIE 133
                ************************************************************

Result          DAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFRK 239
Expected        DAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLPDPQAMFDIEYFRK 193
                ************************************************************

Result          DPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQCHG 299
Expected        DPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLEQVAGIQRIIQCHG 253
                ************************************************************

Result          ------------------------------------------------------------ 324
Expected        SFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEIVFFGENLPEQFHR 313

Result          ------------------------------------------------------------ 327
Expected        AMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPHLHFDVELLGDCDV 373

Result          ------------------------------------------------------------
Expected        IINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTPLHVSEDSSSPERT 433

Result          ------------------------------------------------------------
Expected        SPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESIAEQMENPDLKNVG 493

Result          ------------------------------------------------------------
Expected        SSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYIFHGAEVYSDSEDD 553

Result          -
Expected        V 554
```

E3.9: Protein sequence alignment for SIRT1.9. Alignment of the translated DNA sequencing result after reverse sequencing (TriExDOWN primer) of the expression vector carrying the gene of interest SIRT1.9 with the expected SIRT1 sequence using clustal 2.1. His-tag sequence highlighted in blue.

```
Result          ------------------------------------------------------------
Expected        MAHHHHHHSSGKETAAAKFERQHMDSPPPSGLVPRGSAGSGTIDDDDKCESDEEDRASHA 60


Result          ------------------------------------------------------------
Expected        SSSDWTPRPRIGPYTFVQQHLMIGTDPRTILKDLLPETIPPPELDDMTLWQIVINILSEP 120


Result          ------------------------------------------------------------
Expected        PKRKKRKDINTIEDAVKLLQECKKIIVLTGAGVSVSCGIPDFRSRDGIYARLAVDFPDLP 180


Result          ------------------------------------------------------------
Expected        DPQAMFDIEYFRKDPRPFFKFAKEIYPGQFQPSLCHKFIALSDKEGKLLRNYTQNIDTLE 240


Result          ------------------------------------------------------------
Expected        QVAGIQRIIQCHGSFATASCLICKYKVDCEAVRGDIFNQVVPRCPRCPADEPLAIMKPEI 300


Result          -------PEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPH 53
Expected        VFFGENLPEQFHRAMKYDKDEVDLLIVIGSSLKVRPVALIPSSIPHEVPQILINREPLPH 360
                       **************************************************

Result          LHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTP 113
Expected        LHFDVELLGDCDVIINELCHRLGGEYAKLCCNPVKLSEITEKPPRTQKELAYLSELPPTP 420
                ************************************************************

Result          LHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESI 173
Expected        LHVSEDSSSPERTSPPDSSVIVTLLDQAAKSNDDLDVSESKGCMEEKPQEVQTSRNVESI 480
                ************************************************************

Result          AEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYI 233
Expected        AEQMENPDLKNVGSSTGEKNERTSVAGTVRKCWPNRVAKEQISRRLDGNQYLFLPPNRYI 540
                ************************************************************

Result          FHGAEVYSDSEDDV-PGFSSTISELVDPNSQISARLQVDGTGSKLAAAQLYTRASQPELA 292
Expected        FHGAEVYSDSEDDV---------------------------------------------- 554
                **************

Result          PEDPEDLEHHHHHHHH-VINLRCRLPIRRWW 322
Expected        -------------------------------
```

PUBLICATION

- Ramaswamy, S., Sleiman, M. H., Masuyer, G., Arbez-Gindre, C., Micha-Screttas, M., Calogeropoulou, T., Steele, B. R. and Acharya, K. R. Multivalent galactose based dendrimer recognition by human galectin-7. 2014, FEBS J. [Epub ahead of print] doi: 10.1111/febs.13140.

POSTER PRESENTATIONS

- Ramaswamy, S., Masuyer, G. and Acharya, K. R. Structural study of human galectin-7 and its interaction with novel carbohydrate based dendrimers, 2013, Postgraduate Research Day, University of Bath, U.K.
- Ramaswamy, S., Masuyer, G and Acharya, K.R. Structural study of human galectin-7 and its interaction with novel carbohydrate based dendrimers, 2013, Diffraction data collecton using synchrotron radiation, Helmholtz-Zentrum, Berlin, Germany.
- Ramaswamy, S., Masuyer, G. and Acharya, K. R. Multivalent galactose based dendrimer recognition by human galectin-7, 2014, Southwest structural biology consortium (SWSBC), University of Bath, U.K.
- Ramaswamy, S., Masuyer, G. and Acharya, K. R. Multivalent galactose based dendrimer recognition by human galectin-7, 2014, Postgraduate Research Day, University of Bath, U.K.